THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

VOL. 57, PART 6 November 1945 No. 324

THE ELECTROPHYSIOLOGICAL ANALYSIS OF THE FUNDAMENTAL PROBLEM OF COLOUR RECEPTION

BY RAGNAR GRANIT,
Director of the Nobel Institute for Neurophysiology,
The Royal Caroline Institute, Stockholm

Fourteenth Thomas Young Oration, delivered 29 *June* 1945

§ 1. INTRODUCTION

In a memorable lecture at the Royal Society in 1801 Thomas Young expressed his well-known views on the mechanism of colour reception. This was done almost in passing, and accompanied by the comment that he regarded them as something of a corollary to Newton's great discovery. It was impossible, he said, to conceive of the many spectral colours as being represented by an equally great number of optic nerve fibres of different type. For this reason he suggested that the number is limited. He fixed the number of fibre types at three, but this limitation is less essential than the three main generalizations implied in his thesis: (i) that the analysing mechanism is placed in the peripheral visual apparatus, (ii) that the number of colour-sensitive elements is relatively limited, (iii) that these elements represent widely different regions of the visible spectrum. Certain fundamental aspects of the problems contained in these assumptions can now be investigated with the aid of very direct methods. It has become possible to record the impulses in single fibres of the optic nerve of animals.

It is well known to most physicists that Professor Adrian in this country first showed that all sense organs respond to stimulation by discharging trains of impulses of constant size through their nerve fibres. The electrical nature of the nerve impulse as a conducted wave of negativity has been known since the work of Du Bois Reymond in 1849, but by introducing electrical amplification and by cutting down the nerves from certain receptors until only single fibres were left, Adrian succeeded in demonstrating that the response to variations in stimulus intensity is a variation in frequency of an impulse of constant size in the individual fibres (Adrian, 1932). This result is the basis of modern physiological analysis of all sensory mechanisms. The retina was also included in Adrian's pioneer programme of research (1927), and several important results were obtained. But at that time, in the late twenties, the new technique opened up so many prospects to physiology that the seemingly hopeless and formidable problem of isolating single fibres among the several hundred thousands of the optic nerve was left unsolved.

The American physiologist Hartline (1938) solved this technical problem for the first time by a very painstaking method of microdissection. He used the excised eye of the frog, removed lens and cornea and lifted up fibres from the retina on to his electrode in the region where they converge to form the blind spot. This technique is extremely laborious, and I remember Dr. Hartline telling me once that every eighth experiment succeeded, provided, of course, that the experimenter had acquired sufficient skill to succeed at all. (For a review of Hartline's work, see *J. opt. Soc. Amer.* 1940.)

§ 2. MEASUREMENTS ON ISOLATED FIBRES

Our own microelectrode technique is a great deal simpler. For work on colour reception it is essential that the technique for isolation of fibres should not be too difficult. It must be applicable not only to excised eyes but also to mammalian eyes in the living narcotized or decerebrated animal. The microelectrode is a thin platinum wire, isolated with glass down to the tip. It is applied to the inside of an eye from which cornea and lens have been removed. The microelectrode touches the fibres running along the retina to the blind spot. A schematic picture is shown in figure 1, referring to the rat's retina. This picture also illustrates the fact that several receptors converge towards the single fibre isolated. There are few cones in the rat's retina. One cone and a much greater number of rods make up the convergence unit assumed to be isolated by the microelectrode. Such isolated fibres have very different properties. The records in figure 2 illustrate the simplest type: the isolated unit responds to an increase in intensity from the threshold upwards, with an increasing number of impulses at an increasing rate of discharge. These particular records are from the eye of a guinea pig in which such units form the great majority; it was maximally sensitive to green light around 530 mμ, as shown by figure 3. The ordinate of this curve is the inverse value of the relative quantum intensity necessary to elicit one impulse, as in the uppermost record of figure 2. The animal was light-adapted.

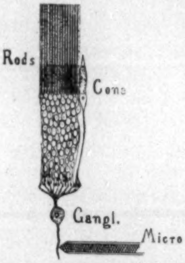


Figure 1. Diagram illustrating microelectrode on optic nerve fibre of rat. Several receptors and bipolars converge to form the unit recorded from. (Granit, *Acta Physiol. Scand.* 2, 1941 a.)

This experiment illustrates the main principles of the microelectrode experiment on colour reception. It is a technique for measuring the absolute threshold to spectral light. In actual practice such experiments are not carried out photographically, but the impulses, or, as we say, the *spikes*, are amplified and led to a loud-speaker. The spectral energy is increased by means of a calibrated wedge until a sharp report is heard in the loud-speaker. Single fibres are not always obtained, and quite often one has to be content with a highly restricted discharge. Since the threshold method always picks up the most sensitive fibre of all, this limitation is less serious than might have been imagined. In my latest experiments with the cat's retina, single fibres were obtained in 60% of the cases.

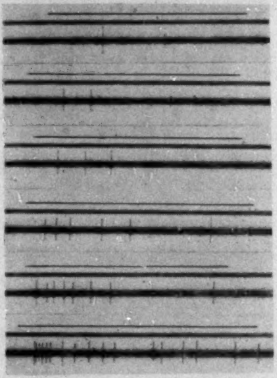


Figure 2. Guinea pig. Isolated unit with maximal sensitivity in 530 mμ responding to different intensities, increasing from record to record downwards. Uppermost line in each record = light signal. Time in 1/50 sec. between this and spike records. Uppermost record at threshold strength, lowermost record at 10·7 times threshold energy. (Granit, *Acta Physiol. Scand.* **3,** 1942.)

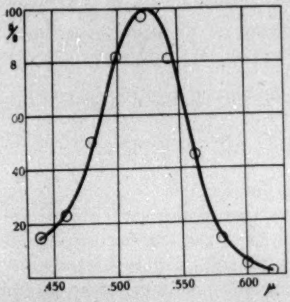


Figure 3. Spectral distribution of sensitivity of spike recorded in figure 2. (Granit, *Acta Physiol. Scand.* **3,** 1942.)

Success depends mainly on the success in making the microelectrode, and to some extent, of course, also on patience and practice. The animal must be quiet, for the slightest movement suffices to shift the microelectrode.

I have said that the fibre shown in figure 2 represented a simple type of response. In the cat's eye, for instance, the fibres respond to an increase of stimulus intensity in a very complex manner. Most fibres discharge to both onset and cessation of illumination. But the relative amount of on- and off-discharges varies a great deal with stimulus intensity. Light may inhibit the on-discharge in a certain intensity range and inhibit the off-discharge in another. The complexity of the response, as soon as the threshold strength has been exceeded, is something stupendous. Hardly two fibres can be said to be exactly alike. An instance is given in figure 4. The figures to the right above each record show the density of the neutral filter put into the beam of a white light of 892 m.c. In this case the fibre is spontaneously active all the time. The on-discharge increases with stimulus intensity (downwards), the off-effect increases first, is inhibited in record 3.09, increases a second time (record 1.38) and is ultimately completely inhibited at maximal intensity (record 0.09). These complex effects are due to interaction in the extremely complex retinal switchboard of neurons and synapses from which the optic nerve fibres take off. Such phenomena must play an important rôle for discrimination and contrast effects, but they do not concern us here. Figure 4 merely serves as a warning against premature and over-optimistic simplification when one is dealing with a wonderful microcosm of a nervous centre such as the vertebrate retina.

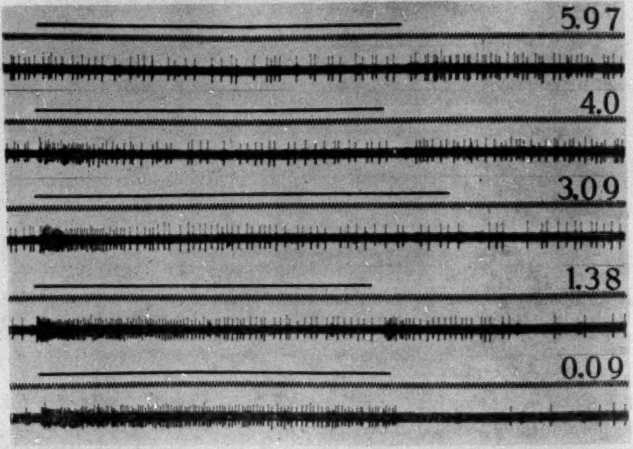


Figure 4. Cat. Isolated spontaneously active element responding to progressively increasing intensities of white light. The numerals to the right show the extinction (densities) of the filter-wedge combination used in front of a light of 892 m.c. Marked as figure 2. (Granit, *J. Physiol.* **103**, 1944.)

§ 3. THRESHOLD MEASUREMENTS FOR COLOUR-RECEPTION ANALYSIS

From the point of view of the analysis of colour reception we must begin with threshold measurements in the manner just described. For this immediate purpose it is immaterial how the discharge is modified by an increase of stimulus intensity. We can always determine the energy necessary for a threshold response. This is the method that has been used in my analysis of the fundamental problem of colour reception, the problem raised with such clarity by the far-sighted genius of Thomas Young.

The results of measurements of the energy necessary for a threshold response at various wave-lengths could be very easily interpreted, were it not for the fact that in most of the laboratory animals both rods and cones converge towards the same fibre, as in figure 1. Now the rods contain the extremely light-sensitive substance *visual purple*, which has maximum absorption around 500 mμ and a quantum yield, as determined by Dartnall, Goodeve and Lythgoe (1936), of 1, so that one quantum of light destroys one molecule of visual purple. Its absorption curve has been fairly well known since the work of Trendelenberg in 1904, and was first accurately determined with up-to-date technique of extraction and photo-chemical analysis by the late Dr. Lythgoe in this country in 1937, and almost simultaneously by several other workers. Schneider, Goodeve and Lythgoe showed in 1939 that the photosensitivity of visual purple agreed with the human scotopic luminosity curve, illustrating the distribution of brightness for a dark-adapted eye to a spectrum of low intensity. (For a full review of this work, see author's summary, Granit, 1946).

It is therefore not surprising that my method reproduces the same curve if the threshold energy in the spectrum is determined for fully dark-adapted eyes of animals containing rods charged with visual purple. Such measurements will be shown later. At the moment this point is taken up in order to emphasize that the visual purple distribution of spectral sensitivity is a serious complication in work with animals. Visual purple can be removed by light-adaptation, but it regenerates during the time the experiment is carried out in the dark, and so tends to raise the sensitivity in the green around 500 mμ. This, of course, is a consequence of the convergence of several receptors of mixed nature towards each optic nerve fibre as well as of the extreme sensitivity to light of visual purple. But when we use our own visual purple mechanism in the dark, the spectrum appears colourless. An increased sensitivity to green around 500 mμ cannot, therefore, in the microelectrode experiment, be ascribed to a hypothetic "green response" of the *cones*. In order to be able to do so we must make certain that we have not been engaged in measuring merely the photosensitivity of visual purple. The experimental work becomes something in the nature of a fight against the absorption curve of visual purple, which tends to cover up the properties of other receptors belonging to the convergence unit which has been isolated by the microelectrode.

§4. THE DOMINATOR AND MODULATORS

In order to illustrate one of the main findings, I shall therefore first draw attention to a curve from the eye of a snake which is lacking visual purple. This animal has a pure cone retina. In figure 5 the curve interrupted by dots refers to the snake (1943 b), the other curve to the frog's eye (1941 c), which has been light-adapted so as to depress the activity of visual purple. The two curves are sufficiently similar to support a conclusion that after light-adaptation the remaining cones of the frog's eye have behaved like the cones of the snake. I shall refer to this broad cone curve with maximum around 560 mμ as the ***dominator***. Those familiar with the problems of vision will immediately notice that the dominator curve represents a distribution of sensitivity which is almost identical with the human daylight luminosity curve for cone vision, the so-called ***photopic luminosity curve***. Provided that an animal's eye possesses a sufficient number of cones, this curve is always obtained, and represents the most common electrophysiological finding. It is lacking in the eyes of rats and guinea pigs (Granit, 1941 a, 1942). The guinea pig's eye is said to lack cones altogether, the rat's eye is said to contain about 1 % cones, apparently too few to be able to give the dominator.

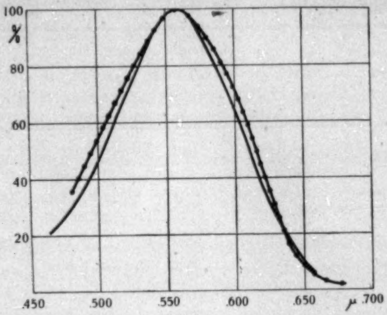


Figure 5. Distribution of sensitivity of dominator element in the retina of snake (line interrupted by dots) and frog (uninterrupted line). (Granit, *Nature*, *Lond.*, **151,** 1943.)

In a sense the dominator can be described as a physiological unit response, but it is very probable that several receptors have combined to form this unit. We must remember that all optic nerve fibre units are convergence-units, as stated above.

When different optic nerve fibres are picked up, in the manner described, the whole experiment is very much dependent upon statistical chance and good luck. These factors favour the dominator, but other types of curve are also sometimes obtained. These are illustrated in figure 6. The curves are from eyes of different animals : rats (1941 a), guinea pigs (1942), frogs (1941 c), snakes (1943 b). First should be noted that they differ from the dominator curve in two important respects : they are much narrower and are spread over a large fraction of the visible spectrum.

Neglecting for the moment the narrow curve with maximum around 500 mμ, which may be described as a narrow visual purple curve, the most striking fact demonstrated by figure 6 is the confinement of the narrow curves to three preferential regions of the spectrum. This is the more remarkable if one considers that they are from different animals. I have called these curves the ***modulator curves***.

The most characteristic red modulator had its maximum at 600 mμ. It was found in rats (dots), much to my surprise, since I took up the rat's rod eye in order to find out what happens to the visual purple distribution of sensitivity when the eye becomes light-adapted. The result was that I found the red modulator remaining, when the green end of the spectrum, occupied by the visual purple absorption, had disappeared below the instrumental threshold of the Hilger-Tutton monochromator used at that time. After some time in the dark the narrow curve with maximum at 500 mμ turned up. Later, during dark-adaptation, it expanded and assumed the shape of the typical visual purple absorption curve. I have since seen the same narrow curve in other rod eyes, but not in the cone eye of the snake. It is difficult to say whether it would play any part in human colour vision or not. Probably not, if central fixation is used. The rat's red modulator may be due to the 1% cones. Histologists, though not all, believe in cones in the rat's eye. There was no dominator in this eye. The red modulator was also seen in the eyes of frogs, in the pure cone eye of the snake, and was indicated as a hump on the green curve in some very rare cases in the guinea pig. In the frog I found some modulators with maximum at 580 mμ. These will be called yellow modulators.

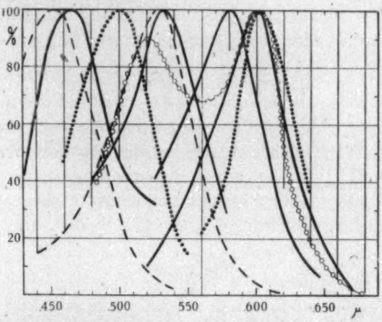


Figure 6. Distribution of sensitivity of modulator elements from eyes of rat (dots), guinea pig (broken line), frog (line in full) and snake (line interrupted by circles). In this figure and in figure 5 spectra of equal quantum intensity. The ordinates on either side of 560 mμ drawn down to indicate dominator values. (Granit, *Nature, Lond.*, **151**, 1943.)

Green modulators are found in all eyes between 520 and 540 mμ, even in the cone eye of the snake, but in this animal it proved impossible to separate them from the red modulator. Blue modulators around 460 mμ were found in the frog's eye and in the pure rod eye of the guinea pig, but not in the cone eye of the snake.

The general biological implications of these findings are clear enough. Nature has not taken the trouble to invent new mechanisms of colour analysis for every new species. It has, as it were, decided upon the principles to be used and then proceeded to use them with what, from our point of view, we might call a greater or lesser degree of perfection. The eyes of the different animals are probably adapted to the life and habits of the species concerned. With regard to modulators, the boundaries between rods and cones do not seem to be very strict. It seems certain that dominators only are found in eyes with a large number of cones. It is possible and even probable that red modulators are cones in the histological sense, but modulators in the short wave-lengths need not be cones.

Blue modulators must be rods in the histological sense, and it is doubtful whether blue modulators can ever be anything else. The experiments suggest that, on the whole, more attention should be paid to photochemistry than to histology. Histological and photochemical definitions as to what is a cone or a rod may disagree. Unfortunately cone photochemistry is still at the stage when it is no exaggeration to state that the less said about it the better. It is, of course, very difficult to extract substances from receptors which to all appearance are practically colourless.

If I may state my own belief it is that the cone substances are modifications due to changes in the bonds linking the visual purple chromophore to its protein body. There is some evidence for this view which cannot be discussed in this connection (see Granit, 1941 b). Some day those working on the photochemistry of visual purple will provide us with important clues to the solution of this riddle. Lythgoe made a beginning when, shortly before his death, he found that visual purple, regenerating from its photoproduct "transient orange", had its absorption curve shifted towards the long wave-lengths. This work was never published.

§ 5. NATURE OF THE DOMINATOR

The microelectrode experiments now described did not suffice to create that feeling of satisfaction with which the completion of a work properly should end. The element of statistical chance involved made them not only extremely tedious but also unsatisfactory; and what was the meaning of the dominator, which was the most common finding in eyes containing a great number of cones? Could it be regarded as being formed by grouped modulators? Those questions became more pressing when it was found that in some cases the dominator of the snake eye had a hump around 600 mμ and that all dominators in the cat's eye had this same hump. Despite this it proved impossible to obtain red modulators in the cat's eye, although this eye is more like the human peripheral retina than any other eye used in my work.

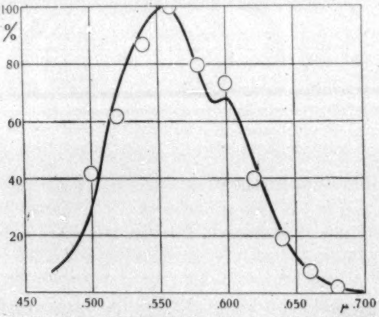


Figure 7. Luminosity curve determined by Wright (*Nature, Lond.*, **151,** 1943) with a small fovea patch of low brightness. The circles illustrate the cat dominator from Granit (*Acta Physiol. Scand.* **5,** 1943 a.) Equal-energy spectrum.

At that time Dr. Wright published in *Nature* (1943) a brief note reporting the results of an experiment in which he had tried to imitate the microelectrode method by measuring the luminosity curve of a very small patch of light. He also found a hump on the foveal luminosity curve in the red around 600 mμ. Figure 7 shows Wright's curve. The circles refer to my own measurements of the cat's dominator. This parallelism between two independent measurements with very different techniques suggested that the dominator might be a composite curve, besides suggesting that the human and the feline mechanism of colour reception cannot be fundamentally different. It therefore became imperative to develop a method by means of which it would be possible to split the dominator, if it could be split. Considering that about 36% of the isolated fibres in the light-adapted cat's eye gave the dominator, a successful splitting of the dominator would minimize the statistical element of chance involved in the process of hunting for simpler units of colour reception. For the new experiments I used Dr. Wright's well-known colorimeter (1934), drawn and constructed for this work by Mr. G. C. Newton.

The principle of the experiment is illustrated in figure 8. The animal is fully dark-adapted so that the highly sensitive visual purple completely dominates and gives the curve *P*. This is based on some 1320 readings, averaged into 179 values, again averaged to give the large circles around the curve. The small black points are Lythgoe's corrected curve for visual purple absorption, which is a little too low in the violet, as shown by Schneider, Goodeve and Lythgoe (1939). A slight effect of the dominator makes my curve a little too high in the yellow-red region.

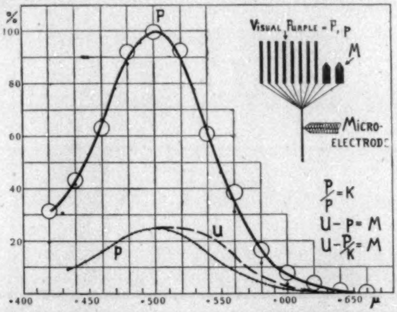


Figure 8. Dark adapted cat. Large circles = average distribution of sensitivity of isolated fibres in the optic nerve. Small black circles = Lythgoe's corrected curve for visual purple absorption from *J. Physiol.* **89**, 1937. Spectrum of equal quantum intensity. As to significance of diagram and symbols, see text. (Granit, *J. Neurophysiol.* **8**, 1945.)

If the eye be adapted to red, blue or green light the curve *P* is reduced to *p*, by proportionate ordinates at all wave-lengths, since *P* represents the homogeneous substance visual purple. Hence

$$P/p = k. \qquad \ldots\ldots(1)$$

But if there are any other colour-sensitive substances (*M*), *preformed* or produced by visual purple, curve *P* does not drop to *p* but to some other curve, *U* (figure 8). The modulators M are given by the difference between U and p, as evident from the diagram, or

$$M = U - p. \qquad \ldots\ldots(2)$$

In this case p is unknown, but the equation can be solved by giving it the form

$$M = U - P/k. \qquad \ldots\ldots(3)$$

P is obtained *before*, U *after*, selective adaptation, and k can be obtained from (1) by finding the spectral region in which selective adaptation has caused the largest drop of sensitivity, since in this region there was no other substance than visual purple left to resist light-adaptation. This region will generally be found in the place where the letter p is inserted in the diagram (figure 8). All quantities of (3) are now known, and the argument can be tested by experiment.

Equation (3) was solved in 34 experiments on the basis of some 4000 observtions collected in 601 points on U-curves. Some 60% of the U-series referred to isolated fibres, the rest to restricted activity. In 29% the equation came out zero ($U = p$), and hence the microelectrode had struck a unit with pure visual purple receptors. In the rest of the series complex curves were obtained.

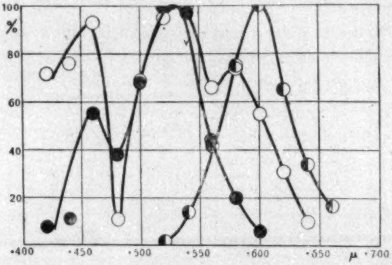


Figure 9. Average effects of selective adaptation of cat's eye, as described in text. ● = red adaptation, ○ = green adaptation, ◐ = blue adaptation. Spectrum of equal energy. (Granit, *J. Neurophysiol.* **8**, 1945.)

Coloured adaptation was used because it was hoped that, for instance, red adaptation would suppress red-sensitive elements and give elements sensitive to other colours a chance to appear. The eye was left in coloured light for some time, this light then instantaneously interrupted, and a test light from the spectrum flashed in, 3 sec. after interruption of the coloured adaptation. This time was chosen in order to give the off-effect time to disappear or to diminish in frequency, so as to make it possible to hear whether the test light caused a fresh discharge or not. When the test had been carried out, the adapting light was again switched on for a while until the eye was ready for a new test with some other spectral wave-length. For adaptation the Ilford spectral filters red, green and blue were used in the main series. A minor number of experiments were carried out with the yellow and violet filters.

For a general survey of the M-curves from equation (3), figure 9 should be studied. This does not show any individual modulators, but merely the general effect of adaptation to red, green and blue light as gross averages of the *M*-curves obtained. The maximum for each individual experiment has always been given the value 100, so that the curves give an idea of the chance for a given point to reach a certain magnitude.

It is seen that blue adaptation gave the most uniform results, so that only curves with maximum in the red were obtained. Red adaptation actually suppressed red modulators, but green and blue modulators made themselves felt. The adaptation to green had a still less selective character. There were humps in both the red, green and blue regions of the spectrum.

Let us now present the same results in a manner which is easier to follow. The individual modulators from the three preferential regions have been picked out and averaged, independently of the kind of coloured adaptation by which they have been obtained. The results are shown in figure 10, in which, also, the dispersion is indicated (outer contours). The narrow red modulators were of two types with maxima respectively at 600 mμ (red) and 580 mμ (yellow), the former type being the more common. Most green modulators overlapped and had maxima at 540 mμ, some at 520 mμ, and two of them were of the type previously described as narrow visual purple curves. These may be lacking in pure cone eyes. These modulators were the ones seen before in different types of retinae, analysed by the earlier "chance" method. The blue region could be better analysed with Wright's colorimeter. Most blue modulators had their maxima at 460 mμ (blue), one at 440 mμ (violet). They were very narrow bands. Errors of measurement increase from the red to the blue end. On account of the steepness of the modulator curves the maxima are fairly well definable.

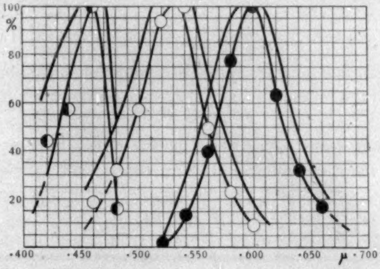


Figure 10. Averaged individual modulators (cat) as obtained by selective adaptation. ●=red modulators, ○=green modulators, ◐=blue modulators. Outer contours indicate dispersion, see text. Spectrum of equal energy. (Granit, *J. Neurophysiol.* **8,** 1945.)

The average modulators of figure 10 are too narrow to add up to a cat dominator. It is clear that impulses also must be delivered up the optic nerve by modulators with "legs" outside the averages. The dispersion gives some idea of the extension of the outer margin of each of the three groups of modulators. In figure 11 the extreme values obtained in these experiments are given. Together with the curves of figure 10, showing the dispersion, they should give some idea of the limits permitted in synthesizing the dominator.

In figure 12 is shown the synthesis of the cat's dominator (Granit, 1943 c) from figure 7. The thick line, drawn through the readings, is the sum of the red and green curves, R and G. In the light of these experiments it seems permissible to maintain that the dominator is a composite curve consisting of modulators. Nevertheless the dominator must be regarded as a *biological unit*. It cannot be neglected in theories of colour vision even though the modulators are the *ultimate physiological units*, the units of first order. It should be emphasized that modulators in the pure state were also obtained without coloured adaptation in some animals.

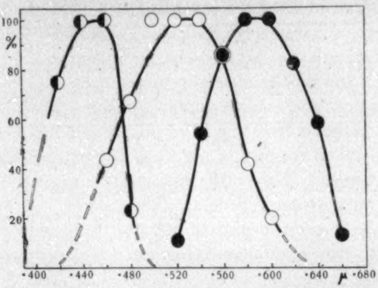


Figure 11. Extreme values obtained for cat modulators by selective adaptation, as described in text. Equal-energy spectrum. (Granit, *J. Neurophysiol.* **8**, 1945.)

Why have modulators never been found in the human eye with sensory methods ? The main reason would seem to be that the analytical unit-field contains too many elements, perhaps up to a hundred thousand nerve fibres. I should also like to draw attention to the fact that the microelectrode technique is a threshold measurement. Wright's and Walters and Wright's (1943) promising results suggest that experiments with small areas and low intensities, perhaps aided by selective adaptation, might open up new possibilities. Some-

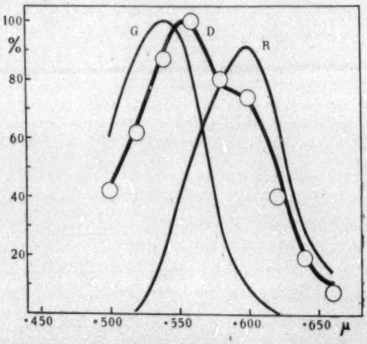


Figure 12. Circles = averages of four photopic dominators of cat. Curve D in heavy lines = synthesis of this dominator by adding modulator areas G and R and plotting their sum in percen of maximum. Equal-energy spectrum. (Granit, *J. Neurophysiol.* **8**, 1945.)

thing might also be gained by studying the peripheral field of vision (see Walters and Wright, 1943).

In figure 13 I have synthesized the human photopic luminosity curve on the basis of the results obtained with the cat's optic nerve. The daylight spectrum or cone spectrum of the human eye agrees very well with the dominator (D) of the electrophysiological experiment. The three preferential regions within which modulators are found are given by the R, G and B curves, forming the three fundamental sensation curves of the Young-Helmholtz trichromatic theory.

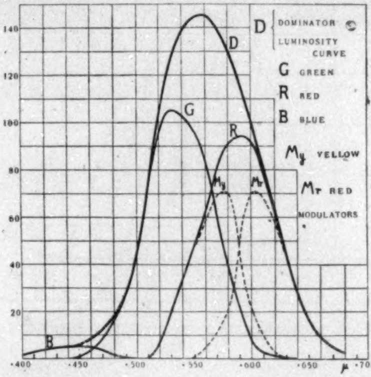


Figure 13. Synthesis of human photopic luminosity curve (D), as determined by Coblentz and Emerson (*Bull. Bur. Stand., Wash.*, no. 303, 1917), on the basis of three fundamental sensation curves, R, G and B. The R-curve indicated to be the sum of two modulators M_y and M_r. Equal-energy spectrum. (Granit, *J. Neurophysiol.* **8**, 1945.)

Their sum is the dominator. For the R-curve I have indicated the two modulators M_r and M_y. It would, no doubt, be possible to combine the modulators in a slightly different manner, for instance an R-curve with a hump at 600 mμ. But the last word on this question must remain with those who are experts on human colour psychophysics. It should be remembered that threshold measurements, such as are used in electrophysiological measurements, record the narrowest curve that can be obtained. It seems possible that the greater the level of intensity, the greater the width of the modulator curve. The retina, after all, is a complex nervous centre with mechanisms of facilitation.

§6. COLOUR VISION

It should be emphasized that the electrophysiological experiments have not provided any evidence for the existence of three fundamental response curves. From the physiological point of view, colour vision must be understood in terms of modulators and dominators. The three fundamental response curves must be regarded as approximations depicting the areas covered by the modulators in the three preferential regions. But it is interesting to note that my schematic response curves agree fairly well with those obtained by Walters (1942) in experiments developing Wright's (1934) method of selective adaptation. Walters' curves are shown in figure 14.

Recent work by Pitt (1944) on the three response curves has not yet become available to me, but from a summary by Stiles (1944) I infer that Walters' results have been confirmed by Pitt by a different method. Pitt's curves extend further down into the blue part of the spectrum.

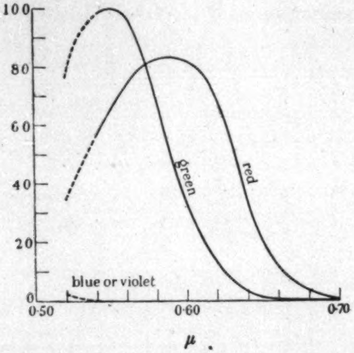


Figure 14. Fundamental sensation curves determined by Walters (*Proc. Roy. Soc.* B, **131,** 1942).

The fineness of hue discrimination would seem to be well explained by the narrow modulator curves. Figure 15 is taken from the work of Wright and Pitt (1934). The ordinates show the shift in wave-length, necessary for discrimination of hue at adjacent wave-lengths, so that minima in the curve signify maxima of discrimination. In the region between 580 and 600 mμ, where the red and the yellow modulators intersect, one can distinguish wave-lengths separated by as little as 1 mμ. From figure 13 it is clear that precisely in this region a slight shift of wave-length suffices to redistribute the nerve fibre signals between

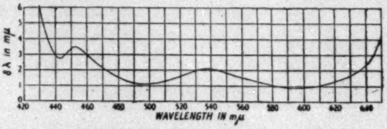


Figure 15. Human hue discrimination according to Wright and Pitt (*Proc. Phys. Soc.* **46,** 1934).

the red and the yellow modulator. The second optimum between 480 and 510 mμ is found in a place where the blue and red modulators meet and intersect with the green ones. Rather remarkable is the third (less marked) optimum between 450 and 440 mμ, which suggests, as did my experiments, that two blue modulators overlap in this region, the blue proper with maximum at 460 and the violet one with maximum at 440 mμ.

§7. THE NATURE OF WHITE AND COLOUR

The existence of the dominator as a physiological unit makes it necessary to consider its sensory equivalent. It is well known that our sensations of light can be divided into two main categories, brightness or luminosity and colour or hue. The luminosity curves of the human eye have several times been referred to. It is well known that corresponding to our two sense-organs in the retina, the rods and the cones, there are two luminosity curves, one for dark-adapted rods charged with visual purple, another for the light-adapted eye and cone vision. These two standard curves are shown in figure 16. The daylight luminosity curve to the right we have identified electrophysiologically as the photopic dominator, the scotopic luminosity curve to the left as a replica of the visual purple absorption curve minus losses of light of short wave-length in the ocular media. If the latter curve is corrected for this and plotted on a quantum basis, it will actually be identical with the curve for visual purple absorption.

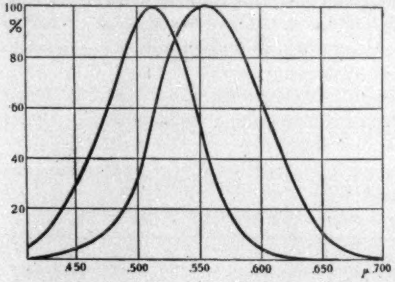


Figure 16. The standard photopic and scotopic luminosity curves of the human eye.

Figure 16 is introduced here merely in order to show that a shift of the maximum need not as such cause an impression of colour. Both curves are luminosity curves illustrating the distribution of brightness. From the physiological point of view it seems reasonable to suggest that optic nerve fibres representing wide spectral areas, such as those of figure 16, cannot cause any other impression than that of brightness or whiteness. For discrimination of wave-length, nature has developed a very different type of mechanism, consisting of narrow modulators in three preferential regions.

According to this view the dominator represents brightness or whiteness. To put it differently, the fundamental visual equivalent to impulse frequency is a neutral white. Colour, if I may say so, is a kind of "local sign", carried by the modulators, which are certain receptors or groups of receptors connected to certain fibres so that from the retina up to the brain the same spatial pattern of organization is maintained. Inasmuch as the modulators contribute to impulse frequency they can also contribute to brightness.

Galambos and Davis (1942) have recently applied the microelectrode technique also to the acoustic nerve. There too it was found that the different acoustic frequency bands were represented by different fibres, rather narrow bands at that. In the ear too the perception of quality or pitch is something related to space, just as in the eye. Loudness, however, is almost certainly a function of total impulse quantity, just as brightness in the eye. Colour and pitch are determined by some kind of local sign.

What makes the eye particularly interesting is that the spectral distribution of brightness is also found to be represented by individual optic nerve fibres. In my opinion this must be regarded as evidence for the view that white has the character of a separate sensation. It has long been known that brightness and colour have maintained a certain amount of independence. This is particularly obvious in defective colour vision. Thus the deuteranope is a red-green blind observer with a normal distribution of luminosity. I do not see how this can be explained by the trichromatic theory, according to which white must be regarded as the sum of the red and green fundamental response curves with a minor contribution from the blue curve. If the red and green curves are lacking, one must expect a grossly abnormal distribution of luminosity instead of the normal one found. This difficulty does not exist for the dominator-modulator concept.

The various forms of colour blindness I regard as defects in the receptors or elsewhere in the paths representing the modulators. But such defects would have to be of an extreme degree in order to be so complete as to remove the dominators, of which, apparently, there is a very great number, considering how easily they can be detected.

There are, too, several forms of acquired colour blindness. They are characterized by loss of perception of certain hues but not of brightness. In certain forms of colour blindness, such as protanopy, there are minor shifts in the luminosity curve. These can be explained by a minor redistribution of the aggregate modulators forming the dominator.

The distribution of saturation in the spectrum is very characteristic. The ends are saturated and at the same time dark, extreme red and extreme blue being examples of such dark and saturated colours. On the dominator-modulator concept this is explained by the fact that at the ends of the spectrum the dominator values are very low, so that the modulators in these regions are responsible for the greater part of the sensory experience. Again, in the yellow region the large contribution from the dominator makes the spectrum appear relatively white and the colour unsaturated.

§ 8. CONCLUSION

Let us now return to Thomas Young, in whose honour this lecture is being delivered. It is a characteristic of those whom we revere as classics of science that their ideas have been formulated with a curious insight into how nature might be expected to behave and so have preserved lasting creative power. What greater tribute could one scientist pay to another's memory than to perform an experiment suggested by his ideas, 140 years after they have been formulated, and come to the conclusion that these ideas were fundamentally correct ? The mechanism of colour reception is organized by the peripheral visual apparatus, the number of colour-sensitive elements is relatively limited, and these elements represent widely different regions of the visible spectrum. Those were Young's three fundamental assumptions. He was right even in assuming three main types of colour-receiving apparatus. These are the three preferential regions within which modulators are found. The electrophysiological work may, indeed, be said to have confirmed the view he gave of the framework of a mechanism of colour reception. Its finished picture looks somewhat different, but the old framework was solid enough and shines through.

It has been a great pleasure for me to have had this opportunity of speaking to this distinguished audience in honour of Thomas Young. His original theme was the same as mine, in the sense that he described a peripheral mechanism of wave-length reception in terms of the properties of nerve fibres. Colour vision is a much wider subject, and there are several in this Society who know much more about it than I do. My theoretical attempts to translate the electrophysiological results into the language of colour psychophysics are legitimate, and I believe them to be correct, but further experience may nevertheless necessitate modifications. I can only hope that I shall not have to make these experiences myself, but that somebody else will try his hand at the optic nerve. I also feel just now that it would be interesting to see for a while what photochemistry and colour psychophysics could do for this field before any further labour is invested in electrophysiological work. For the invitation to present the results of this work to the Physical Society I wish to express my sincerest thanks.

REFERENCES

ADRIAN, E. D., 1932. *Mechanism of Nervous Action* (Oxford: University Press).
ADRIAN, E. D. and MATTHEWS, R., 1927. *J. Physiol.* **63,** 378.
COBLENTZ, W. W. and EMERSON, W. B., 1917. *Bull. Bur. Stand., Wash.*, no. 303.
DARTNALL, H. J. A., GOODEVE, C. F. and LYTHGOE, R. J., 1936. *Proc. Roy. Soc.* A, **156,** 158.
GALAMBOS, R. and DAVIS, H., 1942. *J. Neurophysiol.* **6,** 39.
GRANIT, R., 1941 a. *Acta Physiol. Scand.* **2,** 93.
GRANIT, R., 1941 b. *Acta Physiol. Scand.* **2,** 334.
GRANIT, R., 1941 c. *Acta Physiol. Scand.* **3,** 137.
GRANIT, R., 1942. *Acta Physiol. Scand.* **3,** 318.
GRANIT, R., 1943 a. *Nature, Lond.*, **151,** 11.
GRANIT, R., 1943 b. *Acta Physiol. Scand.* **5,** 108.
GRANIT, R., 1943 c. *Acta Physiol. Scand.* **5,** 219.
GRANIT, R., 1944. *J. Physiol.* **103,** 103.
GRANIT, R., 1945. *J. Neurophysiol.* **8.**
GRANIT, R., 1946. *Sensory Mechanisms of the Retina* (Oxford: University Press. In course of publication).
HARTLINE, H. K., 1938. *Amer. J. Physiol.* **121,** 400.
HARTLINE, H. K., 1940. *J. Opt. Soc. Amer.* **30,** 239.
LYTHGOE, R. J., 1937. *J. Physiol.* **89,** 331.
SCHNEIDER, E. E., GOODEVE, C. F. and LYTHGOE, R. J., 1939. *Proc. Roy. Soc.* A, **170,** 102.
STILES, W. S., 1944. *Proc. Phys. Soc.* **56,** 329.
TRENDELENBERG, W., 1904. *Z. Psychol. Physiol. Sinnesorg.* **37,** 1.
WALTERS, H. V., 1942. *Proc. Roy. Soc.* B, **131,** 27.
WALTERS, H. V. and WRIGHT, W. D., 1943. *Proc. Roy. Soc.* B, **131,** 340.
WRIGHT, W. D., 1934. *Proc. Roy. Soc.* B, **115,** 49,
WRIGHT, W. D., 1943. *Nature, Lond.*, **151,** 726.
WRIGHT, W. D. and PITT, F. H. G., 1934. *Proc. Phys. Soc.* **46,** 459.
YOUNG, T., 1855. *Miscellaneous Works of the late Thomas Young*, ed. by George Peacock, vol. i (London).

THE GEOPHYSICAL ASPECT OF COSMIC RAYS

By ARTURO DUPERIER
London : formerly University of Madrid

Twenty-ninth Guthrie Lecture, delivered 5 July 1945

§1. INTRODUCTION

In my student days I intended, like your founder, to become a chemist. Later, however, physics attracted me in consequence of my interest in thermodynamics, whereby I became acquainted with Professor Guthrie's important discovery of cryohydrates. I could hardly have imagined that the future would hold for me the honour of delivering one of the lectures instituted by this distinguished Society in his memory. This invitation by your Officers and Council is another proof of the kindness shown to me by this country, whose generosity has enabled me to carry on my scientific work.

On the suggestion of Professor P. M. S. Blackett, I undertook in 1939 the study of the variations of cosmic-ray intensity with time. All work carried out in recent years has shown increasingly the close relation of this study with that of terrestrial magnetism and with the physics of the atmosphere. In presenting to you some of the results I have obtained, I shall try to give you an outline of the present state of our knowledge of this geophysical aspect of cosmic rays.

§2. THE TEMPERATURE EFFECT OF COSMIC RAYS

In the last decade or so cosmic rays have been recorded in various parts of the world. The early records, obtained by using ionization chambers, showed an annual variation of the intensity of cosmic rays inverse to that of the temperature near the ground.

It is well known that Blackett (1938) attempted to relate this temperature effect with the instability of the mesons, which form the main part of the penetrating component, and that in this way he explained the decrease of cosmic-ray intensity in warm weather as being due to the greater distance which the mesons have to travel to reach sea-level owing to the greater height of the pressure-level at which they originate. As Blackett pointed out, to test this theory the observed intensity changes would have to be correlated with the mean temperature of the free atmosphere rather than with the temperature near the ground, as had been done until then.

Duperier (1941), using various published cosmic-ray records with the upper-air temperature data for Europe compiled by Wagner and those for the United States compiled by Lennahan, found that the seasonal cosmic-ray variations are related more closely to the average spatial temperature of the atmosphere up to 16 km. than to the temperature near the ground. In particular, it was shown that the lag in the warming of the atmosphere in spring is paralleled by a lag in the diminution of intensity of cosmic rays. But though this was the only way of treating the data then available, it was clear that such a correlation had but slight physical significance, particularly as the observations of atmospheric temperature referred to different intervals of space and time from those of cosmic-rays.

At about the same time, Beardsley (1940) correlated the cosmic-ray intensity at Cheltenham (U.S.A.) and the upper-air temperatures as obtained by radio-sonde balloons released during the period covered by the observations. He found that up to 10 km. the correlations were not appreciably different from those based on ground temperature, but that above this altitude the correlation becomes less close as the height increases, and even changes sign. This would mean that the temperature averaged over a height of 15 km. would be less closely correlated with cosmic-ray intensity.

Later, Hess and Benedetto (1941) and Benedetto, Altmann and Hess (1942) made a similar investigation using the data obtained with a cosmic-ray dual telescope and upper-air temperatures which were observed during the same period. They first correlated cosmic-ray intensity at ground level with temperature values at different levels from the surface up to 12 km., and found that the correlation decreased as the height increased. Because of this result, they were led to a new method of taking mean temperatures. Instead of the ordinary method of integrating the temperature *v.* height curve and dividing by height, they adopted the method of integrating the temperature *v.* pressure curve and dividing by pressure. To the temperature thus defined they gave the name of *mass temperature*. Then they correlated the mean daily cosmic-ray intensity with the mean mass temperature for increasing fractions of the atmosphere, and found that this correlation increases, though only very slightly, as the fraction of the atmosphere becomes greater : from $r = -0{\cdot}71$ for 0·2 to $r = -0{\cdot}75$ for 0·8 of the atmosphere.

Clearly, from these results it would not be possible to say whether the so-called temperature effect of cosmic rays is controlled by the mean temperature of the atmosphere up to a certain level rather than by the temperature near the ground, however the temperature be defined.

But the lack of success in proving the predominant influence of the atmospheric temperature was perhaps not surprising. As a result of the instability of the meson, it is the change in height of the pressure-level at which mesons are formed which gives rise to the fluctuations of the penetrating cosmic rays at the ground, and, as the hypsometric formula shows, the height of any pressure-level depends not only on the temperature, but also on the vertical distribution of temperatures in the air underneath and on the pressure at sea-level. When, in order to find the correlation with temperature, the ground-level meson intensities are first corrected for pressure by using, as is customary, the so-called barometric coefficient, the effect of the pressure change at sea-level on the height is removed, but there still remains the effect of distribution of temperatures to be added to the effect produced by change of mean temperature. Now, since the upper-air observations show that the differences in temperature distribution from day to day are generally far from being negligible, it must be admitted that the simple correlation of daily cosmic-ray changes with mean temperatures cannot lead to convincing results.

On account of this, I have (1944) adopted another method for the analysis of the cosmic-ray records which have been obtained in London in the last few years, by making use of a battery of Geiger-Müller counters registering triple coincidences.

Cosmic rays travelling through the atmosphere lose energy in several ways, the amount depending on the mass of air traversed. In addition, if we accept the principle of the instability of the meson, the penetrating rays are also subject to decay. We may thus establish that the variation of the cosmic-ray intensity at ground level is a function of, firstly, the variation of the mass of air (represented by the barograph reading), and, secondly, the change in height of the pressure-level at which mesons are generated, assuming that there is only one meson-generating layer. Now should this be correct, there must clearly be a correlation between the variations in intensity of cosmic rays and those in height of this pressure-level. But as we do not know where mesons originate we have to put to the test a few pressure-levels for which meteorological information is available.

For the computation the hourly numbers of cosmic rays and the hourly barograph readings were averaged in groups of 24 hours. Of these daily mean data, those corresponding to periods of great geomagnetic disturbance were discarded as well as those corresponding to days on which the sounding balloon at the meteorological station failed to reach an altitude of 16 km., so that the correlations of cosmic rays with the heights of different pressure-levels up to this limit will be entirely comparable.

The following table gives the values of the partial correlations for the pressure-levels which have been chosen.

Table 1

Pressure-level	Partial correlation coefficient
75 mm. Hg (16·1 km.)	−0·67
113 ,, (13·5 ,,)	−0·54
188 ,, (10·3 ,,)	−0·32
375 ,, (5·5 ,,)	−0·30

The gradual increase of the correlation with height shown by the table proves that it is not the temperature near the ground, or at least that this is not the only factor, which is responsible for the temperature effect of cosmic rays, since the influence of this temperature on the height of a certain pressure-level decreases with the mean height of this pressure-level. On the other hand, the significance of the value −0·67 for the correlation between cosmic rays at a constant pressure and the heights of the 75-mm. pressure-level can be taken as confirming the view that a part of the variation of the intensity of the cosmic radiation at ground-level may be explained by spontaneous disintegration of mesons in the atmosphere.

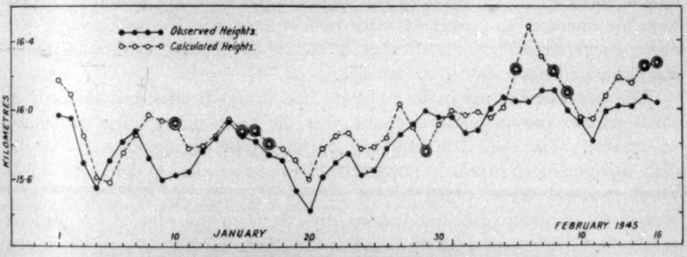
In order to check these results by means of other experiments, we may assume that mesons originate at this pressure and write the regression equation

$$C - C_m = \mu(B - B_m) + \mu'(H - H_m),$$

where C, B and H are the number of cosmic rays, barograph readings and heights respectively, and subscript m refers to mean values. Clearly μ represents the *true* absorption coefficient in air, and μ' the mean rate of decay of mesons. By solving this equation in the usual manner, we find (Duperier, 1944) for the first coefficient $\mu = 2{\cdot}28\%$ per cm. Hg, a value which is practically the same as the value that can be deduced from the measurements by Ehmert (1937) of the absorption curve in water.

For the mean rate of decay of mesons we obtain $\mu' = 5{\cdot}4\%$ per km. This coefficient is the reciprocal of the so-called *mean range* of mesons before disintegration, and is related to the lifetime of mesons when at rest, τ_0, by the very well known equation $\mu' = \dfrac{1}{L} = \dfrac{cM}{E\tau_0}$, where M is the rest mass and E the energy of the mesons. If we assume that the mesons with which we are concerned have a mean energy of 3×10^9 ev., and take M equal to 200 times the mass of an electron, we have $\tau_0 = 2{\cdot}1 \times 10^{-6}$ sec., which is of the same order of magnitude as the values found in other modern experiments, which generally vary between 2 and 3×10^{-6} sec.

These results suggest the possibility of applying the cosmic-ray records at the ground to foretell the daily mean height of the 75-mm. pressure-level and thence, by means of the hypsometric formula, the mean temperature of the air up to about 16 km. The calculation of the height would be made as follows.



DAILY MEAN HEIGHTS OF THE 7·5 CM. HG. PRESSURE-LEVEL
JANUARY 1ST. – FEBRUARY 16TH. 1945.

Figure 1.

We should first reduce the daily average cosmic-ray intensity to a constant pressure by making use of the *true* absorption coefficient. Then the percentage departures of these averages from the mean value divided by the mean rate of decay of the mesons would give the fluctuations in height of the pressure-level.

Figure 1 enables us to compare the heights as calculated in this manner with the actual heights as observed with sounding balloons for the period 1 January to 16 February 1945. The computation of the above coefficients had been made long before. The open double circles indicate the days on which a geomagnetic disturbance was recorded at Abinger. As the diagram shows, the correspondence between the two curves is, on the whole, striking, but on days of magnetic activity, and on the immediately preceding or following days, the discrepancies are both considerable and irregular. We should therefore run the risk of making a grave error in trying to foretell the mean temperature of the atmosphere overhead from the variations of cosmic-ray intensity at the ground.

However, as we shall see later, there is now some evidence that these cosmic-ray disturbances at periods of magnetic activity are of a world-wide character, having probably the same value at points on the same geomagnetic latitude. If this is so, the difference between the cosmic-ray intensities at two stations in a narrow zone of latitude, by eliminating the superimposed disturbance, would give a measure of the difference in temperature of the atmospheres over the two stations. Clearly, if the temperature of the atmosphere at a certain place was known by observation, it would be possible to determine the temperature of the atmosphere at all points in a narrow zone of latitude within which cosmic-ray records were being made. For this, it would, of course, be necessary for all the recorders to be sufficiently stable to give consistent results.

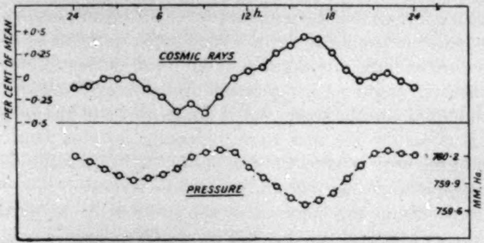
§ 3. DIURNAL VARIATION

Besides the annual variation in cosmic rays which results from the effect of the thermal expansion and contraction of the atmosphere on the number of mesons reaching ground level, records show an appreciable change in cosmic radiation with the time of day. This solar diurnal variation in intensity has been measured by many observers, generally by using lead-shielded ionization chambers, and the results agreed in giving a maximum intensity around noon with an amplitude of about 0·2%. It has been also found, in particular by Thompson (1933) from the extensive measurements made by Compton and Turner during a year on the Pacific Ocean between latitudes 50° N. and 40° S., that both the amplitude and phase are independent of the latitude.

The fact that the maximum intensity has always been found to occur at about midday proved, without doubt, that the temperature effect cannot be responsible. The change in height of the meson-producing layer due to the daily variation in temperature, if appreciable, would give rise to exactly the opposite fluctuation. Hess (1936) pointed out the possibility of the existence of a component of cosmic radiation coming directly from the sun. Later Vallarta and Godart (1939) gave a theory based on the effect on primary cosmic rays of a magnetic field at the surface of the sun of the required magnitude to prevent cosmic-ray particles of less than 2×10^9 ev. energy from reaching the earth. By assuming that the positively charged particles of the incoming radiation were in greater number than the negative ones, these workers found a diurnal variation at latitudes higher than 40° with the maximum occurring during the early afternoon, in agreement with the observed curve. If the ratio of positive to negative particles is taken as $\frac{4}{3}$, the amplitude also agrees with the measurements. For latitudes lower than 40°, Vallarta and Godart found that the effect of the sun's magnetic field became negligible. More recently, Jánossy and Lockett (1941), after analysing the effect on primary cosmic rays of different hypothetical heliomagnetic fields, came to a similar conclusion—that the diurnal variation at 45° latitude can be accounted for by assuming a magnetic field at the surface of the sun. They found also that the amplitude of the diurnal variation must be least in autumn and spring and greatest in summer and winter.

When the harmonic analysis of the observed curves is made, however, the results of different investigators are not always in agreement. It is found that there is a significant 24-hour wave with an amplitude of about 0·2%, which is approximately the same in all cases, though the time of the maximum can vary by as much as eight hours from one station to another. In the same way, the possible existence of a 12-hour wave with an amplitude of about 30% that of the first harmonic is not excluded by some investigators. Forbush (1937) thought that the variations of the 24-hour wave might be due to variations in gamma radiation with locality. According to him, Wait has obtained results indicating the possibility of a daily change of ionization in the air due to gamma radiation, with a maximum at about noon. In passing through the 12-cm. lead shield the gamma rays would produce ionization inside the chamber which would give rise to an additional wave having an amplitude one-third that observed in cosmic rays.

I should like now to present the results, not yet published, that I have obtained at the Imperial College concerning the diurnal variation of cosmic-ray intensity by using the data for the period May 1941 to April 1944. The data for days of great geomagnetic activity have been discarded. Figure 2 shows the average hourly numbers of threefold coincidences, or cosmic-ray particles, as percentages of the mean. As the total number of coincidences recorded for each hour of the days within three years amounts to about 24 millions, the statistical fluctuation has the very small value of 0·02%. The box which contains the set of Geiger-Müller counters is kept at constant temperature by thermostatic control. The electric clock of the recorder has been compared with a chronometer at intervals of a few days over a period of several months, and no systematic difference between day-time and night-time has been found. The variation shown by the diagram can therefore be considered as physically real. Figure 2 also illustrates the average daily variation of pressure in mm. Hg obtained from my hourly readings during the same days of the three-year period.



DIURNAL VARIATIONS OF COSMIC-RAY INTENSITY AND ATMOSPHERIC PRESSURE AT LONDON.

Figure 2.

The harmonic analysis of both curves gives, for the 24-hour wave,

$$\text{Cosmic rays } 0{\cdot}25 \cos\frac{360^\circ}{24}(t-17{\cdot}3)\ \%\ ; \quad \text{Pressure } 0{\cdot}14 \cos\frac{360^\circ}{24}(t-4{\cdot}0)\ \text{mm.},$$

where t is the solar time in hours.

Owing to the additive property of harmonic coefficients, to correct for air mass absorption we subtract from the vector representing the observed 24-hour wave in cosmic-ray intensity that corresponding to the same harmonic in pressure after multiplying the latter by the absorption coefficient −0·23% per mm. But there still remains another correction to be made for pressure. We have to remove the effect on cosmic-ray intensity at ground-level of the change in height of the meson-producing layer which is in phase with the barometric variation. If we assume that the relative pressure change, $\delta p/p$, at the height of 16 km. is the same as that at the ground, the amplitude of the vertical motion of the layer would be only 1·5 m., and the cosmic-ray intensity at ground-level would change by less than 0·01%. If, as is believed, the 24-hour variation of the barometer, which has its minimum at 4 p.m., is mainly the result of lateral flow of air from warmer towards cooler regions of the atmosphere, there is no reason to expect greater relative pressure change at 16 km. than at the ground. Quite probably, therefore, the 24-hour wave in cosmic-ray intensity after correcting completely for pressure would be $0{\cdot}22 \cos \frac{360^\circ}{24} (t - 17{\cdot}3)\%$.

With regard to the effect of daily temperature change, the examination of upper-air data obtained in England in the last few years by sending up sounding balloons every six hours shows that the height of the 75-mm. pressure-level is, on the average, appreciably greater at noon than at other hours. This is to be expected, since it is known that the temperature of the lower stratosphere is mainly controlled by radiation from the earth. We must clearly have a 24-hour wave due to temperature with the minimum at about noon in the observed variation of cosmic-ray intensity. I am not in a position to-day to give you the exact average value of the amplitude of this wave, but it seems to be of the order of 0·1% or rather less. It is easy to see that after correcting for this temperature effect, the maximum of the final 24-hour wave in cosmic-ray intensity will occur earlier in the afternoon and be greater than the observed maximum.

In order to determine the influence of the position of the earth in its orbit, the data have been averaged for the four-monthly seasons, December solstice, June solstice, Equinoxes. Table 2 contains the results obtained for the amplitudes and times of maximum of the 24-hour wave corrected for pressure.

Table 2. 24-hour wave in cosmic-ray intensity corrected for pressure

	Amplitude	Time of maximum
	%	h
Dec. solstice	0·22	15·9
Equinoxes	0·18	16·7
June solstice	0·32	18·5

As the table shows, the time of the maximum is gradually delayed as we pass rom cooler to warmer months and the effect of temperature becomes greater. Had this effect been eliminated, it is quite probable that the time of the maximum would have been the same and would have been nearer to noon for the three seasons, in keeping with the theoretical result of Vallarta and Godart. Another interesting feature is that the amplitude at the solstices is greater than at the equinoxes, in qualitative agreement with the calculations of Jánossy and Lockett, already quoted.

As, moreover, the gamma radiation is not effective in producing coincidences in my apparatus, these results for the 24-hour wave, after correction for meteorological influences, may be taken as supporting the view that one part of the diurnal variation of cosmic-ray intensity is due to the effect on the primary particles of a magnetic field at the surface of the sun. The possibility, however, of explaining the 24-hour wave in cosmic-ray intensity as an effect of the earth's magnetic field, as suggested by Gunn (1932), cannot be excluded.

The harmonic analysis of the triennial curve reveals the existence also of a physically significant 12-hour wave. This wave is given below, together with the semi-diurnal barometric oscillation from my barograph readings.

$$\text{Cosmic rays} \qquad 0{\cdot}18 \cos \frac{360^\circ}{12}(t-15{\cdot}2)\% \, ;$$

$$\text{Pressure} \qquad 0{\cdot}22 \cos \frac{360^\circ}{12}(t-10{\cdot}2) \text{ mm.}$$

As we see, they are within one hour of being opposite in phase.

As before, to remove from the cosmic-ray wave the total effect of pressure we should have to know the amplitude of the vertical motion of the meson-producing layer, which is in phase with the barometric oscillation. If it is assumed that the relative pressure change at the 16 km. level is the same as at the ground, that amplitude would not be greater than 2·5 m., but if we take into account that the solar semi-diurnal oscillation of the atmosphere is partly tidal in character, then the vertical displacement with which we are concerned might prove to be greater.

I may perhaps remind you that this oscillation, which appears all over the world with a contant phase and a regular variation of amplitude with latitude, seems to be excited tidally and thermally by the sun. Its large magnitude is considered to be due, as Kelvin suggested, to magnification by resonance, the atmosphere having a natural mode of free oscillation of period very near 12 solar hours, as shown by Pekeris (1937), when a certain vertical distribution of air temperature is assumed.

In developing this theory, Pekeris finds that the vertical motion of air particles due to the solar barometric oscillation is in phase at all levels with the change of pressure at the ground. Later on, he calculated the magnitude of the vertical displacement of an air particle in the ionosphere at the equator (see Appleton and Weekes, 1939) and found that at 125 km. it should be of the order of 1·45 km. The equivalent displacement for an air particle in the ionosphere over London should be, taking into account the variation in barometric oscillation with latitude, 323 m.

Experimental evidence on the possible magnitude of the increase of the vertical motion of air particles with height is provided by the results of Appleton and Weekes (1939) relating to variation with time of the equivalent height of the Kennelly-Heaviside layer. They found a lunar tide in this layer which is in phase with the lunar barometric oscillation at the ground, as found first by Chapman, and 7000 times greater, the amplitude of the tide in the Kennelly-Heaviside layer being of the order of 1 km.

It seems reasonable therefore to attribute to the layer at 16 km. a vertical displacement of amplitude greater than 2·5 m. In order to form an idea of the effect of a greater motion of this layer on the cosmic-ray intensity at the ground, let us assign to it an amplitude of only 22 m. in contrast to the 323 m. found by

Pekeris for the ionosphere. Then, by applying our value for the rate of decay of mesons, and allowing for the air mass absorption, we should find that the semi-diurnal variation of cosmic-ray intensity would be entirely accounted for by the motion of the meson-producing layer due to the barometric oscillation. Assuming this is so, if the 12-hour wave in cosmic rays were merely an effect of the 12-hour wave in atmospheric pressure, we should expect the seasonal changes of the two semi-diurnal variations to follow each other closely. Now this is in fact what my results reveal. Table 3 enables us to compare the amplitudes and times of second maximum of both waves for the solstices and the equinoxes.

Table 3

	Amplitude		Time of second max.	
	P	*C.R.*	*P*	*C.R.*
	mm.	%	h	h
Solstices	0·21	0·16	10·2	15·4
Equinoxes	0·25	0·22	10·1	14·7

As the table shows, the amplitudes, within the limits of accuracy of the observations, are closely correlated and the phases remain constant.

It may be that mesons are generated at a greater height than 16 km., as suggested by the results of the experiments of Schein, Jesse and Wollan (1941). If this is so, it would be an additional reason for concluding that the semi-diurnal variation of cosmic-ray intensity at ground-level can represent the barometric oscillation in the region of the atmosphere where mesons are formed. In this connexion it would be interesting to see whether the moon has an appreciable influence on cosmic rays.

As the barometric oscillation is greater at lower latitudes, we must expect the magnitude of the semi-diurnal change of cosmic radiation to vary in the same direction, though the longer lifetime of mesons as we approach the equator tends to offset the effect.

Unfortunately I cannot compare my results with those of other workers, since it has been the practice to analyse the data after correcting them for pressure by using the so-called barometric coefficient.

§4. WORLD-WIDE CHANGES

In the last decade many investigators have reported certain effects on cosmic-ray intensity during great magnetic storms. The most notable of these effects is a decrease of a few per cent in the cosmic-ray intensity which is concomitant with a decrease in the horizontal component of the earth's magnetic field.

The magnetic storm of 1 March 1942 provides a most remarkable example on account of the exceptional magnitude of the decrease and other features of the storm effect on cosmic rays. At that time cosmic-ray recorders were in operation in various widely separated parts of the world, and a comparison of the results obtained reveals the world-wide character of the phenomenon. Figure 3 illustrates the changes of the bi-hourly mean number of cosmic rays expressed in percentages of the pre-storm value, as observed in London. The data, reduced to constant atmospheric pressure, were obtained with the coincidence arrangement mentioned previously. The diagram indicates that the regular decrease began shortly before the onset of the magnetic storm, and reached the unusually high value of about 11%. After the first rapid recovery of about half the initial drop, which nearly coincided with the end of the storm, after midnight, the recovery was very slow, probably owing to the various geomagnetic disturbances which took place during the following days. Other features worth noticing are the increase on 28 February and the much greater one on 7 March. Increases as large as that of 7 March seem to be quite exceptional.

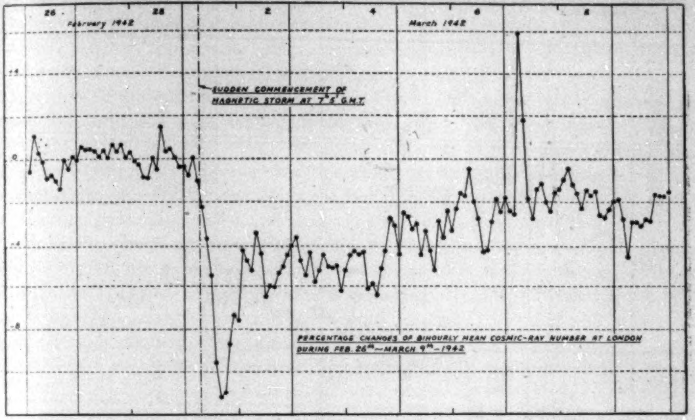


Figure 3.

This diagram differs from that published at the time (1942) in that bi-hourly means are used now to facilitate comparison with the diagrams published later.

Figure 4 shows the diagram, as given in a note by Lange and Forbush (1942), of bi-hourly changes in cosmic-ray intensity at Godhavn (Greenland) in geomagnetic latitude 80° N., Cheltenham (Maryland) in geomagnetic latitude 50° N., and Huancayo (Peru) in geomagnetic latitude 0·6° S. during the same magnetic storm. At each of these three stations the data were obtained with Compton-Bennett meters protected by a 10-cm. lead shield. [In the same note, another diagram is given representing the bi-hourly changes, during the same storm, at Christchurch (New Zealand), and this reveals precisely the same features.] The very striking similarity of the simultaneous changes in cosmic-ray intensity at the three stations and London proves the world-wide character of the phenomenon beyond doubt. Apart from the smaller value of the increase on 28 February in London, the only outstanding discrepancy is the non-occurrence at Huancayo of any of the increases on 28 February and 7 March, which took place simultaneously at all the other stations. This particular result seems to suggest that the new particles responsible for both increases were unable to reach the magnetic equator, and, therefore, their energy was within the energy range which is sensitive to the magnetic field of the earth. Instead, the magnitude of the drop, within the limits of accuracy of bi-hourly means, is roughly the same at all stations, though owing to the fact that the altitude of Huancayo is 3350 metres and that of the other stations is less than 100 it is not possible to say whether any variation of the drop with latitude exists.

The solar phenomena which preceded the onset of this storm were in some respects exceptional. As reported by Newton (1942), a great sunspot crossed the solar disk between 22 February and 7 March. The maximum area of the spot was 2000 millionths of the sun's hemisphere. A very extensive and brilliant eruption was observed over the umbra of the great sunspot nineteen and a half

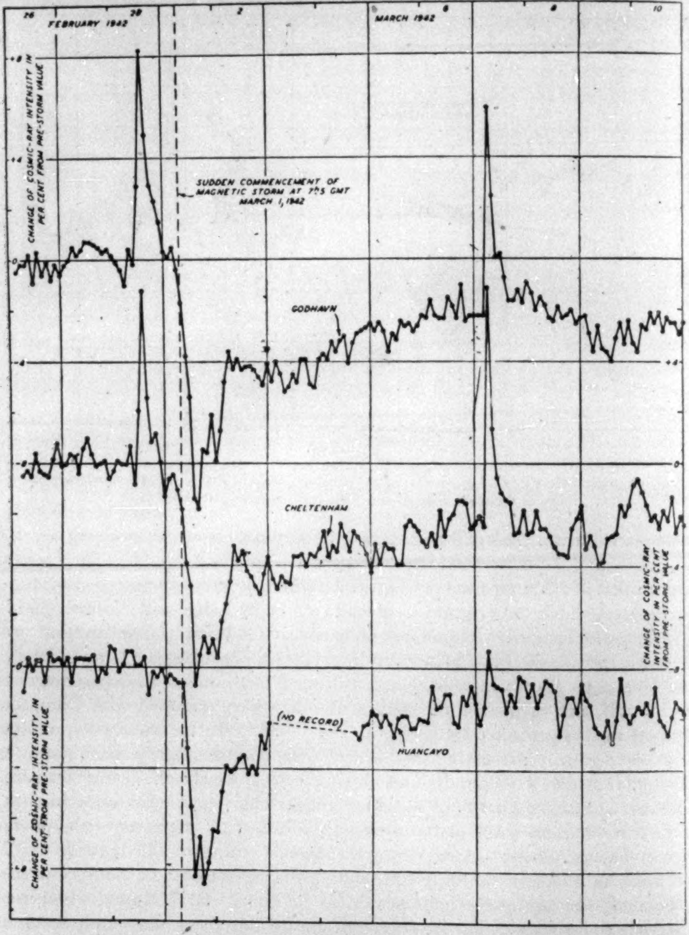


Figure 4.

hours before the beginning of the storm. This eruption appears to have been of exceptionally long duration.

Judging, however, from the magnitude of the range in the horizontal magnetic force, H, of 512γ, the magnetic storm was not as intense as some others which have been accompanied by a much less violent disturbance in cosmic-ray intensity. To give an example, the range in H of the great magnetic storm of 4 July 1941 was, from the Abinger traces, about twice the range of the storm of 1 March 1942. Nevertheless, no appreciable effect on cosmic-ray intensity was shown by the records obtained in London for the 40-hour period following the abrupt commencement of the storm at $3^{\text{h}}\,42^{\text{m}}$ a.m. of 4 July. Only at the end of this period, when the storm was over, did a decrease begin, and even then it was very small (less than 3%). But as another, though moderate, storm occurred a few hours later, on 6 July, it is possible that this small decrease was associated with the second disturbance rather than with the first one, and if so, the conclusion would be that the great storm of 4 July had no effect on cosmic radiation. Forbush (1938) has reported the case of the magnetic storm of 21 August 1937, during which no effect was detectable in the cosmic-ray records of Cheltenham (U.S.A.), Teoloyucan (Mexico) or Huancayo (Peru). However, all observers agree that the ratio of the relative change in cosmic-ray intensity to that of the horizontal component, while being always positive, differs in different storms.

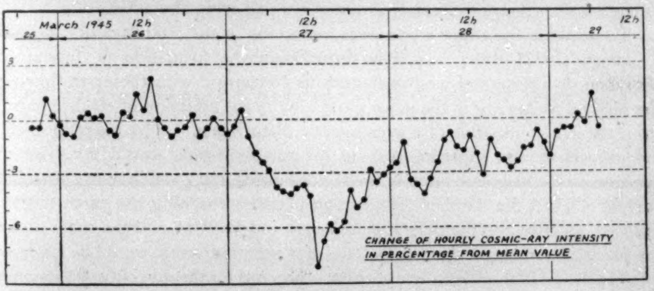


Figure 5.

When smaller changes in cosmic-ray intensity which cannot be explained as being due to meteorological influences are considered, the observations made in London show that they are also more or or less closely associated with variations of the geomagnetic field. These fluctuations generally occur simultaneously with, or within a short time of, the fluctuations in the magnetic field of the earth, but as in the case of magnetic storms, the two variations may be of quite different magnitudes. The magnitude of the changes in the daily means of cosmic-ray intensity associated with these minor geomagnetic disturbances only rarely proves to be greater than 2% of the mean value. In this regard, however, a decrease of a very exceptional order of magnitude was observed in London on 27 March 1945. Figure 5 shows, in percentages of the mean value, the hourly change of cosmic-ray intensity, reduced to constant pressure, for a period of a few days around 27 March 1945. Within the limits of accuracy of hourly observations, the decrease of about 8% indicated by the diagram, which was not associated with a magnetic storm, is, as far as I know, unprecedented in all records of cosmic rays. According to information given by the Royal Observatory, nothing abnormal was observed either in the magnetic activity or in the solar phenomena during that period, although a comparatively slight magnetic disturbance was recorded at Abinger, chiefly on the day preceding that of the decrease in cosmic-ray intensity. A sunspot, but only of the moderate area of 750 millionth of the sun's hemisphere, crossed the central meridian on the 27th after midnight. The complexity of these phenomena is evident.

These fluctuations in cosmic-ray intensity associated with magnetic or solar activity should be considered as taking place simultaneously all over the world. Striking evidence that the world-wide character of variations in cosmic radiation is not confined to periods of great magnetic storms has been given by Forbush (1939). He made the analysis of the data obtained at five widely separated stations and found that after eliminating a seasonal wave of a 12-monthly period which appeared at each of the stations except at Huancayo, there existed a very close correspondence between the remaining curves for all stations.

In addition, we have the finding, first reported by Hess and Graziadei (1936), of a 27-day periodicity in cosmic-ray intensity associated with the mean period of the sun's rotation relative to the earth. These results, therefore, can be taken as giving strong evidence that all appreciable changes at ground-level in cosmic radiation other than meteorological are of a world-wide character.

To account for these world-wide variations, some investigators, following Chapman (1937), think they may result from the formation or fundamental alteration of a system of westward currents concentric with the earth flowing in the high ionosphere or in the outer space, due to the emission of electric particles from the sun at the time of a geomagnetic disturbance. The magnetic field of this current-system superimposed on the earth's dipole would give rise to a decrease of the magnetic force in the regions inside and to an increase outside the current. Then the number of incoming particles reaching the earth might be reduced, according to the theory of Lemaitre and Vallarta, and the main phase of the phenomenon, particularly in the case of magnetic storms, would be explained.

If by means of this theory it is possible to account for the diversity of phenomena observed, we shall have to wait, as Forbush has pointed out, until the difficult problem of determining the paths of cosmic-ray particles in the field of the earth's dipole and the system of westward currents has been solved.

There is no doubt that the relation between cosmic rays and magnetic activity presents a very complicated problem. But I am convinced we shall not be discouraged by the difficulties involved in its solution. For when we have solved them we shall know more not only about cosmic rays, but also about the complex phenomena of geomagnetism.

REFERENCES

Appleton and Weeks, 1939. *Proc. Roy. Soc.* A, **171**, 171.
Beardsley, 1940. *Phys. Rev.* **57**, 336.
Benedetto, Altmann and Hess, 1942. *Phys. Rev.* **61**, 266.
Blackett, 1938. *Phys. Rev.* **54**, 973.
Chapman, 1937. *Nature, Lond.*, **140**, 423.
Duperier, 1941. *Proc. Roy. Soc.* A, **177**, 204.
Duperier, 1942. *Nature, Lond.*, **149**, 579.

DUPERIER, 1944. *Terr. Mag.* **49,** 1.
DUPERIER, 1944. *Nature, Lond.,* **153,** 529.
EHMERT, 1937. *Z. Phys.* **106,** 751.
FORBUSH, 1937. *Terr. Mag.* **42,** 16 ; 1938. **43,** 203.
FORBUSH, 1939. *Rev. Mod. Phys.* **11,** 168.
GUNN, 1932. *Phys. Rev.* **41,** 683.
HESS, 1936. *Terr. Mag.* **41,** 9 and 345.
HESS and BENEDETTO, 1941. *Phys. Rev.* **60,** 610.
JÁNOSSY and LOCKETT, 1941. *Proc. Roy. Soc.* A, **178,** 52.
LANGE and FORBUSH, 1942. *Terr. Mag.* **47,** 331.
NEWTON, 1942. *The Observatory,* **64,** 260
PEKERIS, 1937. *Proc. Roy. Soc.* A, **158,** 650.
SCHEIN, JESSE and WOLLAN, 1941. *Phys. Rev.* **59,** 615.
THOMPSON, 1938. *Phys. Rev.* **54,** 93.
VALLARTA and GODART, 1939. *Rev. Mod. Phys.* **11,** 180.

THE THERMAL EXPANSION OF GRAPHITE FROM 15° C. TO 800° C.: PART I. EXPERIMENTAL

By J. B. NELSON AND D. P. RILEY,
Cavendish Laboratory, Cambridge
(Now at British Coal Utilisation Research Association)

MS. received 23 *March* 1945

ABSTRACT. The variation with temperature of the a and c unit-cell dimensions of hexagonal Ceylon graphite has been measured over the temperature range 15°–800° C. by the X-ray powder method. At 14°·6 C.,

$$a=2{\cdot}4562\pm0{\cdot}0001\text{ kx.}\quad c=6{\cdot}6943\pm0{\cdot}0007\text{ kx.}$$

The carbon-carbon bond length,

$$\text{C}-\text{C}=1{\cdot}4210\pm0{\cdot}0001\text{ A.}$$

The a dimension shows a slight contraction up to about 400° C., a small expansion occurring above this temperature. The thermal expansion in the c direction is large ; the average value for α over the temperature range is $28{\cdot}3\times10^{-6}$. The complex nature of the expansion in both directions is discussed qualitatively.

§1. INTRODUCTION

A STUDY of the thermal expansion of graphite is of interest for two main reasons. The first is the need for extending our knowledge of the solid state in the field of anisotropic crystals ; the second is the importance of precise information concerning the effect of heat on elementary carbon in connection with the technological processes of carbonization and coking. The X-ray method is the only suitable one for an investigation of this character. Other methods, the results of which are summarized in §2, are not able to do more than measure the *mean* linear expansion of the crystal. Such information is of little significance thermodynamically, the essence of the problem being the *anisotropy* of the structure. The work reported here shows how valuable a tool X-ray diffraction methods can be in measuring the effect of temperature on such structures.

§ 2. PREVIOUS WORK

Previous X-ray work comparable to ours has been carried out by Backhurst (1923) in the course of an investigation of the variation in intensity of X-ray reflections with temperature. He measured with the ionisation spectrometer and molybdenum radiation the shift of the 0004 reflection (called by him 0001) on heating from 17° to 870° C. In this way an average value over this range of $26{\cdot}7\times10^{-6}$ was derived for $\alpha_{\parallel}$, the coefficient of linear expansion in the *c* direction.* This value is somewhat lower than the average value of $28{\cdot}3\times10^{-6}$ found by us. He did not measure the coefficient in the *a* direction, $\alpha_{\perp}$, but surmised that a contraction and not an expansion might occur.

Hirata (1931) has investigated the behaviour of carbon in an incandescent arc at a temperature of about 4000° C. Even at this temperature, no measurable expansion in the *a* direction occurred, but by observing the shift of the 0002 reflection (with copper radiation) he derived these average values for $\alpha_{\parallel}$ over this large temperature range :

$\alpha_{\parallel}=39{\cdot}1\times10^{-6}$ assuming a temperature of 4200° C. by measurements of surface brightness.

$\alpha_{\parallel}=45{\cdot}1\times10^{-6}$ assuming an average temperature of 3700° C.

The thermal expansion of graphite has also been measured by various other methods, and the following values have been recorded :

	Expansion coefficient	Temperature (°C.)
Fizeau (1869)	$7{\cdot}86\times10^{-6}$	40°
	$7{\cdot}96\times10^{-6}$	50°
Muraoka (1881)	$3{\cdot}8\ \times10^{-6}$	26° to 302°
Day and Sosman (1912)	$0{\cdot}55\times10^{-6}+3{\cdot}2\times10^{-9}t$	0° to 1800°
Hidnert and Sweeney (1927)	$2{\cdot}7\ \times10^{-6}$	20° to 600°
	$3{\cdot}7\ \times10^{-6}$	
Erfling (1939)	$2{\cdot}34\times10^{-6}$	−195° to −183°
	$6{\cdot}695\times10^{-6}$	+20° to +40°
Dewar (1902)	$7{\cdot}33\times10^{-6}$ (vol. coeff.)	−190° to +17°

In no case was a single crystal of graphite used. As the specimens probably consisted of aggregations of small graphite crystals variously oriented, the values can only refer to the mean coefficient of linear expansion $\bar{\alpha}$. In a hexagonal crystal

$$\bar{\alpha}=\frac{2}{3}\alpha_{\perp}+\frac{1}{3}\alpha_{\parallel}$$

and, using our data, $\bar{\alpha}\approx8\times10^{-6}$ at 20° C. Several of the values are much too low and probably refer to the expansion of the aggregated solid as a whole rather than to the expansion of the individual crystals. It is interesting that Fizeau, who was the first to measure the thermal expansion of graphite in the course of his classical researches, was also responsible for the best value of $\bar{\alpha}$.

---

* The symbols $\alpha_{\parallel}$ and $\alpha_{\perp}$ refer to the coefficients of linear expansion parallel and perpendicular to the hexagonal axis.

The work of Erfling requires special comment. Employing an optical method, he attempted to measure $\alpha_{\perp}$ by taking care to orient the crystals in the specimen. In spite of his precautions Erfling must have been measuring the mean coefficient of expansion. His value of $6{\cdot}695 \times 10^{-6}$ at 20° to 40° c. is probably a low value for $\bar{\alpha}$; it is much too large to be $\alpha_{\perp}$, which we have found to be a small negative quantity over this range.

§ 3. APPARATUS

The camera used was the high-temperature 19 cm.-diameter powder-camera described by Wilson (1941). The camera angle θ_k for this instrument has been carefully determined by Wilson and Lipson (1941). The use of this camera in the determination of accurate unit-cell dimensions had been previously investigated (Nelson and Riley, 1945). The method of measuring the temperature of the specimen was essentially that of Wilson. We found, however, that the melting point of aluminium filings did not provide a satisfactory high-temperature calibration as the aluminium reacted with the silica capillary in which it was contained. Antimony (Hilger H.S.), with a melting point of 630°·5 c., was used in its place. It can be said that the error in temperature measurement did not exceed ± 1° c. at the highest temperature investigated and was certainly less than this at the lower temperatures.

§ 4. PREPARATION OF THE SPECIMEN

Some difficulty was encountered in obtaining a specimen of graphite which combined high purity with the ability to give the sharp diffraction lines necessary for accurate measurements.

An examination of powder photographs of available graphites ("Kish", Acheson, Mexican, Bavarian, Travancore, Ceylon) showed that the Ceylon specimen gave the sharpest lines, although the purity was not of the highest (4% ash). It was felt, however, that purification by the usual chemical treatments (boiling with NaOH, H_2SO_4 or HF solution) was to be avoided because of the ease with which graphite forms complexes which show an appreciably altered *c* dimension. The Ceylon graphite was therefore chosen in spite of its low purity.

After reducing the material to pass a 300-mesh B.S.S. sieve, several silica capillary tubes of bore approximately equal to 0·05 cm. and a wall thickness of 0·003 cm. were filled. It was thought advisable to remove traces of adsorbed gases by heating to about 850° c. during the evacuation of the filled capillaries. However, as soon as even gentle heating was applied to them, the graphite powder was completely swept out, presumably by gases released during the heating. This made it necessary to heat the powdered graphite in bulk in a large-bore (1·5 cm.) continuously evacuated silica tube. Heating for 30 minutes at approximately 850° c. was found to be an ample treatment for completely removing the gas. Photographs of the graphite taken in an ordinary powder camera showed that the only extra lines discernible, besides those due to a small amount of the rhombohedral modification of graphite, were those of Fe_3O_4 in the heat-treated material and hydrated Fe_2O_3 in the untreated. It seems fairly certain, therefore, that the evolution of gas is connected with the reaction

$$3Fe_2O_3 + C \rightarrow 2Fe_3O_4 + CO$$

together with the evolution of water vapour.

No trouble was encountered in evacuating, heating and sealing off the capillaries after heat-treatment. The sealed-off capillary with the smallest bore and thinnest, most uniform wall thickness was selected for the exposures, which were made in order of descending temperature. Its bore, as measured by a microscope ocular micrometer, was 0·048 cm. and the wall thickness 0·0025 cm.

Subsequent precision measurements of both the *a* and *c* dimensions of the heat-treated and untreated graphites at room temperature were in excellent agreement. It would appear, therefore, that the impurities present (chiefly oxide of iron) have no influence on the measurements as they simply form a mixture with the graphite itself.

§ 5. THE X-RAY PHOTOGRAPHS

X-ray photographs were taken at the following temperatures: 14°·6, 150°, 300°, 450°, 600°, 700°, 800° c.

The temperature fluctuation during the exposure was greater the higher the temperature, but in no case did it exceed ± 0°·5c. The photographs were taken with unfiltered manganese radiation with exposures of about 200 ma. hrs.

The high-angle portions of the x-ray photographs are reproduced in figure 1. Even a qualitative comparison leads to some interesting conclusions. The *hki*0 reflections ($11\bar{2}0$ and $20\bar{2}0$) do not move with change of temperature, while the 0006 doublet moves towards the low-angle end of the film quite markedly with increasing temperature. It can also be observed with the *hkil* reflections that as the *l* index increases ($11\bar{2}2$, $11\bar{2}4$, $10\bar{1}5$), the rate of line shift with temperature also increases. The thermal expansion of the lattice is clearly anisotropic, the expansion occurring along the *c* axis. It should be noted that, at 15° c., the 553 doublet of Fe_3O_4 is practically coincident with the $11\bar{2}4$ β line of graphite. Resolution is effected at higher temperatures due to the different thermal expansions.

Another effect which is clearly visible concerns the relative intensities of the diffraction lines. The relative intensity of the 0006 reflections compared with the intensities of the *hki*0 reflections ($11\bar{2}0$, $20\bar{2}0$) shows a progressive diminution with temperature. This fading of the 0006 line is due to the large amplitude of thermal vibrations perpendicular to the layers, and obviously is in agreement with the pronounced anisotropy of thermal expansion. The $10\bar{1}5$ reflection shows a similar effect because of the high *l* index.

There are also nine extra lines due to the Fe_3O_4 previously mentioned, although only one (553 doublet) can be seen in figure 1. The identity of this compound was established by comparison of the interplanar spacings and intensities of the lines observed with that of pure Fe_3O_4. To obtain the true cell dimensions, the calculated apparent length of the cubic unit cell was plotted against the extrapolation function $\frac{\cos^2\theta}{\sin\theta} + \frac{\cos^2\theta}{\theta}$ for the four strongest lines on both the room

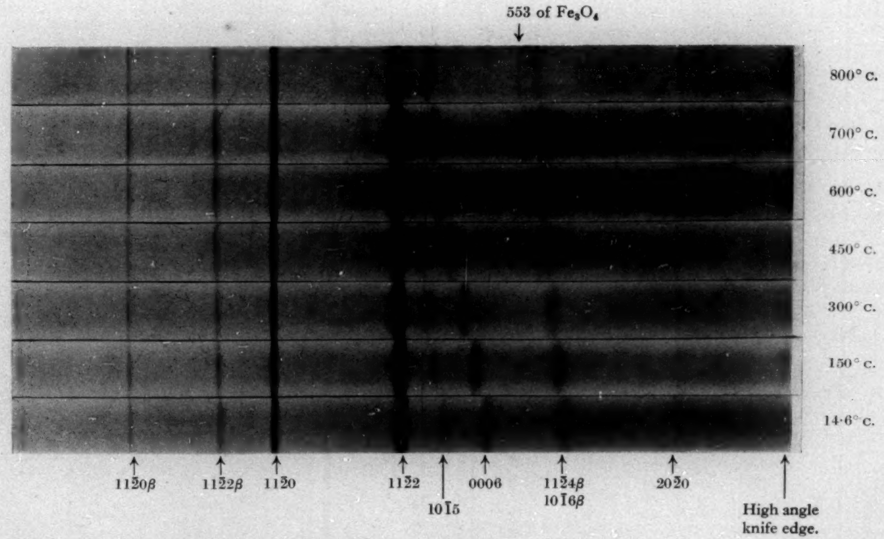

553 of Fe_3O_4
800° c.
700° c.
600° c.
450° c.
300° c.
150° c.
14·6° c.
$11\bar{2}0\beta$
$11\bar{2}2\beta$
$11\bar{2}0$
$11\bar{2}2$
$10\bar{1}5$
0006
$11\bar{2}4\beta$
$10\bar{1}6\beta$
$20\bar{2}0$
High angle
knife edge.

Figure 1. X-ray photographs of graphite at various temperatures.

temperature and 800° c. films. The results are as follows :

Temperature (°c.)	*a* (kx.)
15	8·378
800	8·475
20 (pure Fe_3O_4, Wyckoff and Crittenden 1925)	$8{\cdot}37 \pm 0{\cdot}01$

Mean linear coefficient of thermal expansion, $14{\cdot}7 \times 10^{-6}$ (range 15° to 800° c.)

§6. MEASUREMENT OF UNIT-CELL DIMENSIONS AND DERIVATION OF EXPANSION COEFFICIENTS

The best radiation to use (manganese) gives the $20\bar{2}0$ α-doublet at $\theta = 80°$ to 81° and the 0006 α-doublet at $\theta \approx 70°$ at room temperature. These angles, and particularly the latter one, are not as high as is desirable in precision measurements, but it is impossible to effect any improvement in this case by change of x-ray wave-length. Hence, particular attention had to be paid to the extrapolation procedure adopted.

Work described elsewhere (Nelson and Riley, 1945) has shown that a reliable linear extrapolation is obtained down to $\theta \approx 30°$ if the unit cell dimension under consideration is plotted against $\frac{\cos^2\theta}{\sin\theta} + \frac{\cos^2\theta}{\theta}$. The apparent *c* dimensions given by the 0004 ($\theta = 38°{\cdot}9$ at room temperature) and 0006 reflections were therefore plotted against this function. A linear extrapolation to $f(\theta) = 0$ gave the *c* value quoted for each temperature. As the extrapolation lies through only two points, the accuracy claimed is mainly dependent on the suitability of the function of θ used in the plot. We have assigned an error of about one part in 10,000 to the *c* values derived in this way. The 0004 lines were accompanied by a weak satellite line on all the films. This feature was shown to arise from a preferred orientation of the graphite crystals in the specimen. The effect of this on the accuracy of the determination of the *c* dimension has already been discussed (Nelson and Riley, 1945).

In order to obtain the best value for the *a* dimensions at each temperature, the apparent *a* values derived from the following lines were plotted against $\frac{\cos^2\theta}{\sin\theta} + \frac{\cos^2\theta}{\theta}$:

$20\bar{2}0\ \alpha_1, \alpha_2$.
$11\bar{2}0\ \alpha_1, \alpha_2, \beta$.
$10\bar{1}0\ \alpha, \beta$.

At room temperature, the θ values for the $10\bar{1}0$ α and β lines were roughly 29°·6 and 26°·6 respectively. We did not assign them much weight in the extrapolation, but they serve to confirm its general trend. The error assigned to any given *a* dimension is approximately one part in 25,000. The refraction correction for *a* or *c* is negligible in comparison with the probable experimental error.

The values derived for the *c* dimensions at the various temperatures are given in table 1. A plot of *c* against temperature gives a smooth curve deviating only slightly from linearity. An expression in t and t^2 was adopted to represent this curve, and the results were analyzed by the method of least squares. The best-fitting relation was found to be

$$c=6{\cdot}6915_4+180{\cdot}70\times10^{-6}t+12{\cdot}63\times10^{-9}t^2, \qquad \ldots\ldots(1)$$

where t is expressed in °c. and c in kx.

Table 1. c dimensions of graphite at various temperatures, in kx

Temperature (°c.)	Observed	Calculated	Difference (calc.-obs.)
14·6	6·6943	$6{\cdot}6941_8$	$-1{\cdot}2\times10^{-4}$
150	6·7185	$6{\cdot}7189_3$	$+4{\cdot}3$
300	6·7474	$6{\cdot}7468_8$	$-5{\cdot}2$
450	6·7752	$6{\cdot}7754_1$	$+2{\cdot}1$
600	6·8046	$6{\cdot}8045_1$	$-0{\cdot}9$
700	6·8240	$6{\cdot}8242_2$	$+2{\cdot}2$
800	6·8443	$6{\cdot}8441_8$	$-1{\cdot}2$

The values of c calculated from this equation are also given in the table. The deviations between the calculated and observed values are too large for equation (1) to be considered a satisfactory expression. The mean deviation is $2{\cdot}4\times10^{-4}$, and c is therefore expressed with an average accuracy of only one part in 2800, whereas the observed values themselves are accurate to about one part in 10,000. That the assumption of a parabolic expression for c is not justified is confirmed by examining the variation with temperature of the expansion coefficient. Equation (1) leads to the following linear expression for the coefficient $\alpha_{\|}=\dfrac{1}{c}\cdot\dfrac{\partial c}{\partial t}$:

$$\alpha_{\|}=27{\cdot}00\times10^{-6}+3{\cdot}05\times10^{-9}t \qquad \ldots\ldots(2)$$

The straight line of equation (2) is compared with values of $\alpha_{\|}$ derived directly from the experimental data in figure 2.

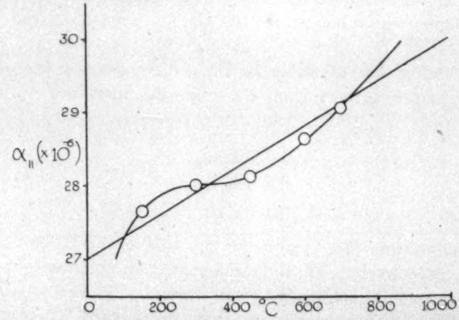


Figure 2. Comparison of observed $\alpha_{\|}$ values with the best linear representation.

These values of $\alpha_{\|}$ were derived in the following way. From the values c_1, c_2, c_3 of the c dimension at neighbouring temperatures t_1, t_2, t_3, two extreme values for $\frac{\Delta c}{\Delta t}$ at the middle temperature t_2 were obtained. The mean of these, divided by c_2, was taken as the best value of $\alpha_{\parallel} = \frac{1}{c} \cdot \frac{\Delta c}{\Delta t}$ at the temperature t_2. Owing to the large values of Δt involved, this derivation of $\alpha_{\parallel}$ cannot be very accurate and, in particular, it is impossible to judge its course at either end of the temperature range.

It would appear from figure 2, however, that the temperature variation of $\alpha_{\parallel}$ is of a complex form and is only approximately represented by a linear expression in t. A more exact representation is derived on theoretical grounds in Part II. The average value for $\alpha_{\parallel}$ over the temperature range is $28 \cdot 3 \times 10^{-6}$.

The measurements of the *a* dimension at the various temperatures showed that a shrinkage followed by an expansion occurred. In order to check this, the films were measured independently by each of us, and the best extrapolated *a* values are given in table 2.

Table 2

Temperature (°c.)	*a* dimensions in kx.	
	Measurement A	Measurement B
14·6	2·45626	2·45617
150	2·45575	2·45564
300	2·45570	2·45563
450	2·45577	2·45571
600	2·45574	2·45574
700	2·45607	2·45602
800	2·45613	2·45619

It will be seen that the agreement between the two sets of measurements is good. Two further checks were carried out in order to confirm the general nature of the effect.

(*a*) In measurement A the θ values for $20\bar{2}0$ α_1 at 15° and 300° c. are 80°·419 and 80°·499 respectively. The difference is therefore 0°·080 which corresponds to a line shift of 0·27 mm. This distance lies outside the limits of error in observing line positions (±0·10 mm.). Therefore we conclude that there is a real change in angle and a real diminution in the *a* dimension.

(*b*) The measured θ values for the $20\bar{2}0$ and $11\bar{2}0$ lines were plotted against temperature for both sets of measurements, A and B. A general rise and then fall in θ was observed, in accordance with the shrinkage and subsequent expansion of *a*. This showed that the extrapolations could not be systematically in error.

Rough values of the thermal contraction and thermal expansion coefficients can be derived from table 2, and are:

Temperature range	$\alpha_{\perp}$
0° to 150° c.	$-1 \cdot 5 \times 10^{-6}$
ca. 400° c.	0
600° to 800° c.	$+0 \cdot 9 \times 10^{-6}$

The *a* dimension first contracts and then expands. It reaches its lowest value of approximately 2·4557 kx. at a temperature of about 400° c.

The unit cell dimensions at 15° c. are

$$a=2{\cdot}4562 \pm 0{\cdot}0001 \text{ kx} ; \quad c=6{\cdot}6943 \pm 0{\cdot}0007 \text{ kx}.$$

and the carbon-carbon bond length, expressed in A., is

$$\text{C—C}=1{\cdot}4210 \pm 0{\cdot}0001.$$

We found no evidence that the rhombohedral modification of graphite (Lipson and Stokes, 1942) has a different thermal expansion from that of the hexagonal variety. Our measurements, however, refer specifically to the latter.

§7. DISCUSSION

A qualitative explanation of the above results will be given first, in which the treatment of Grüneisen and Goens (1924) (see also Roberts (1940)) of the thermal expansion of hexagonal crystals will be followed.

Graphite is a perfect example of a layer structure. Each layer consists of a very large number of carbon atoms covalently linked to form what may be considered to be a huge polynuclear aromatic macromolecule. Each layer is planar, or very nearly so. The carbon atoms within a layer are very strongly linked, the C—C distance being about 1·42 A. The bonding between adjacent layers, however, is weak, as is indicated by the large interlayer spacing of about 3·35 A. It is probable that the interlayer forces are mainly of a van der Waals type, although they may not be entirely so. The graphite structure is therefore one of great anisotropy, and this will be reflected in all its directional physical properties.

The elastic deformability parallel to the principal Z axis will be much larger than that perpendicular to it owing to the tighter interatomic bonding within the layers. This may be written in terms of the elastic moduli as $s_{33} \gg s_{11}$.

Under the influence of temperature and of zero-point energy the atoms will vibrate, and the limiting frequency ν_{m_z} for vibrations parallel to the axis will be much smaller than for those perpendicular to it, ν_{m_x}. The solid will therefore possess two characteristic temperatures defined by these two limiting frequencies.

$$\Theta_z = \beta\nu_{m_z},$$

$$\Theta_x = \beta\nu_{m_x},$$

where

$$\beta = \frac{\boldsymbol{h}}{\boldsymbol{k}}$$

$$\boldsymbol{h} = \text{Planck's constant}, \quad \boldsymbol{k} = \text{Boltzmann's constant}$$

and

$$\Theta_x > \Theta_z.$$

At low temperatures, quantum theory requires nearly all the thermal energy to be absorbed in oscillations parallel to the axis because of the low value of Θ_z as compared with Θ_x. These oscillations will, therefore, have much the greater amplitude and the thermal expansion parallel to the axis, $\alpha_{\parallel}$, will be appreciable, whereas that perpendicular to the axis, $\alpha_{\perp}$, will be negligible.

Thus the lattice is effectively being stretched in one direction only and a lateral contraction (Poisson contraction), proportional to the modulus s_{13}, will accompany the stretching. At low temperatures, the lateral contraction will be greater than the small thermal expansion within the layers, and $\alpha_\perp$ will therefore be negative. As the temperature increases, more energy will go into vibrations perpendicular to the axis, until a temperature T_m is reached when the thermal expansion exactly balances the lateral contraction. At this temperature $\alpha_\perp = 0$. Above this temperature the thermal expansion within the layers will increasingly outweigh the lateral contraction, and $\alpha_\perp$ will be positive and increase with temperature. When this is the case an elastic contraction parallel to the axis ($\propto s_{13}$) will be contributed, which will have the effect of diminishing the rate of increase of $\alpha_\parallel$ and at a high enough temperature may cause $\alpha_\parallel$ to diminish after passing through a maximum. The apparent flattening of the curve for graphite between 250° and 450° c. (figure 2) can be explained in this way.

Grüneisen and Goens (1924) observed with zinc and cadmium similar effects to those for graphite but at lower temperatures. Zinc and cadmium are hexagonal layer structures but their anisotropy is not so pronounced as that of graphite. The data for the metal are compared with those for graphite in the following table.

	T_m	Θ_x	Θ_z
Cadmium *ca*	45° K.	214	160
Zinc ,,	80° K.	320	200
Graphite ,,	660° K.	2280	760

The values of Θ for graphite are those given by Magnus (1923).

It will be noted that Θ_x for graphite is very much higher than for Zn or Cd as is the ratio Θ_x/Θ_z. It would therefore be expected that T_m also would be very much higher. Grüneisen and Goens themselves suggest, when briefly discussing the implications of their work in the case of graphite, that this substance would have $\alpha_\perp$ negative up to higher temperatures than does Zn or Cd.

§8. ACKNOWLEDGMENTS

We wish to thank Professor Sir Lawrence Bragg, F.R.S., for giving us hospitality and working facilities in the Cavendish Laboratory. The work constitutes part of a research programme of the British Coal Utilisation Research Association on the structure and properties of carbon and carbon derivatives, and we are grateful to Dr. D. H. Bangham for his encouragement and interest. Dr. A. R. Stokes gave us access to his collection of graphite samples and x-ray photographs, and we are indebted to him and to Dr. H. Lipson for several helpful discussions.

REFERENCES

Backhurst, I., 1923. *Proc. Roy. Soc.* A, **102**, 340.
Day, A. L. and Sosman, R. B., 1912 a. *J. Acad. Sci., Wash.*, June 19, 284; 1912 b. *Industr. Engng. Chem.* **4**, 490.
Dewar, J., 1902. *Proc. Roy. Soc.* A, **70**, 237.
Erfling, H. D., 1939. *Ann. Phys., Lpz.*, **34**, 136.
Fizeau, H., 1869. *C.R. Acad. Sci., Paris*, **68**, 1125; *Ann. Phys., Lpz.*, **138**, 26.
Grüneisen, E. and Goens, E., 1924. *Z. Phys.* **29**, 141.

Hidnert, P. and Sweeney, W. T., 1927 a. *Phys. Rev.* **29**, 371 ; 1927 b. *Tech. Pap. Bur. Stand., Wash.*, **21**, 223.
Hirata, M., 1931. *Sci. Papers Inst. Phys. Chem. Res., Tokyo*, **15**, 219.
Lipson, H. and Stokes, A. R., 1942. *Proc. Roy. Soc.* A, **181**, 101.
Magnus, A., 1923. *Ann. Phys., Lpz.*, **70**, 303.
Muraoka, H., 1881. *Ann. Phys., Lpz.*, **13**, 307.
Nelson, J. B. and Riley, D. P., 1945. *Proc. Phys. Soc.* **57**, 160.
Roberts, J. K., 1940. *Heat and Thermodynamics*, 4th ed. (Cambridge: University Press).
Wilson, A. J. C., 1941. *Proc. Phys. Soc.* **53**, 235.
Wilson, A. J. C. and Lipson, H., 1941. *Proc. Phys. Soc.* **53**, 245.
Wyckoff, R. W. G. and Crittenden, E. D., 1925. *J. Amer. Chem. Soc.* **47**, 2876.

THE THERMAL EXPANSION OF GRAPHITE: PART II. THEORETICAL

By D. P. RILEY,
Cavendish Laboratory, Cambridge
(Now at British Coal Utilisation Research Association)

MS. received 23 *March* 1945

ABSTRACT. A theory of the thermal expansion of hexagonal crystals is derived, and shown to account quantitatively for the experimental data on graphite. Certain of the elastic moduli for graphite are estimated, and are at 18° c. approximately:

$$s_{11}+s_{12}= 1{\cdot}8\times10^{-13}\ \text{cm./dyne}$$
$$s_{13}=-4{\cdot}3\times10^{-13}\quad ,,$$
$$s_{33}= 58{\cdot}5\times10^{-13}\quad ,,$$

Hence the two linear compressibilities are, at 18° c., approximately:

$$\kappa_{\perp}=-2{\cdot}5\times10^{-13}\ \text{cm}^2/\text{dyne}\ ;\qquad \kappa_{\parallel}=50\times10^{-13}\ \text{cm}^2/\text{dyne}.$$

§1. INTRODUCTION

Grüneisen and Goens (1924) derived a theory of the thermal expansion of hexagonal crystals at low temperatures which enabled them to correlate the expansion data for zinc and cadmium with the elastic moduli and heat capacities. In order to achieve this for the experimental results for graphite described in Part I, a theory which is not restricted to low temperatures must be used. The object of this paper is to derive such a theory and apply it to the data for graphite. The approach is somewhat more straightforward than that of Grüneisen and Goens. In particular, the heat-capacity quantities used are more simply defined and are easily obtainable from the experimental data usually available.

§2. THERMAL EXPANSION OF HEXAGONAL CRYSTALS

In expressing thermal expansion, the basic thermodynamic relationships are:

$$\left(\frac{\partial V}{\partial T}\right)_P=-\left(\frac{\partial S}{\partial P}\right)_T$$

or

$$\left(\frac{\partial^2 V}{\partial T^2}\right)_P=-\frac{1}{T}\left(\frac{\partial C_p}{\partial P}\right)_T,$$

where V is the volume of the system, S the entropy, T the absolute temperature, and C_p the heat capacity at constant pressure.

Integration gives

$$V\alpha = \left(\frac{\partial V}{\partial T}\right)_P = -\int_0^T \left(\frac{\partial C_p}{\partial P}\right)_T \frac{dT}{T},$$

where α is the volume coefficient of expansion, the integration being performed at constant pressure.

Analogous expressions hold for the two linear expansion coefficients in a hexagonal crystal (Voigt, 1910). In this case, the x and y axes are equivalent and

$$\alpha_\perp = \alpha_x = \alpha_y = \frac{1}{l_x}\left(\frac{\partial l_x}{\partial T}\right)_P; \quad \alpha_\| = \alpha_z = \frac{1}{l_z}\left(\frac{\partial l_z}{\partial T}\right)_P,$$

where l_x and l_z are linear dimensions in the x and z directions, and $\alpha_\perp$ and $\alpha_\|$ are the linear expansion coefficients perpendicular and parallel to the hexagonal axis.

In the case of a monatomic solid, if v is the volume and S the entropy of a gram-atom, then

$$v\alpha_x = -\left(\frac{\partial S}{\partial X_x}\right)_T; \quad v\alpha_z = -\left(\frac{\partial S}{\partial Z_z}\right)_T,$$

where X_x is the pressure along x on a plane perpendicular to x, and where Z_z has a similar meaning. X_x and Z_z are sometimes called "thermal pressures" and are positive if the interior of the crystal is considered. Thus, since

$$S = \int_0^T \frac{C_p}{T}\,dT,$$

then

$$v\alpha_x = -\int_0^T \left(\frac{\partial C_p}{\partial X_x}\right)_T \frac{dT}{T}; \quad v\alpha_z = -\int_0^T \left(\frac{\partial C_p}{\partial Z_z}\right)_T \frac{dT}{T}, \qquad \ldots\ldots(1)$$

where C_p is the heat capacity at constant pressure per gram-atom. $\dfrac{\partial C_p}{\partial X_x}$ is now written as

$$\frac{\partial \log l_x}{\partial X_x}\frac{\partial C_p}{\partial \log l_x} + \frac{\partial \log l_y}{\partial X_x}\frac{\partial C_p}{\partial \log l_y} + \frac{\partial \log l_z}{\partial X_x}\frac{\partial C_p}{\partial \log l_z},$$

and the elastic moduli * introduced by using the relations

$$\frac{\partial \log l_x}{\partial X_x} = -s_{11}; \quad \frac{\partial \log l_y}{\partial X_x} = -s_{12}; \quad \frac{\partial \log l_z}{\partial X_x} = -s_{13}.$$

Since for the hexagonal system

$$\frac{\partial C_p}{\partial \log l_y} = \frac{\partial C_p}{\partial \log l_x},$$

---

* The elastic moduli s_{11}, s_{13}, s_{33} are fourth-order tensors and can be written in full as s_{1111}, s_{1122}, s_{1133}, s_{3333}.

For a stress along X,

s_{11} is the modulus giving the strain in the X direction,
s_{12} ,, ,, ,, ,, Y ,,
s_{13} ,, ,, ,, ,, Z ,,

For a stress along Z,

s_{13} is the modulus giving the strain in the X direction,
s_{33} ,, ,, ,, ,, Z ,,

In the discussion, s_{pq} are assumed to be independent of temperature.

we have
$$-\frac{\partial C_p}{\partial X_x} = (s_{11}+s_{12})\frac{\partial C_p}{\partial \log l_x} + s_{13}\frac{\partial C_p}{\partial \log l_z}.$$

Similarly,
$$-\frac{\partial C_p}{\partial Z_z} = 2s_{13}\frac{\partial C_p}{\partial \log l_x} + s_{33}\frac{\partial C_p}{\partial \log l_z}.$$

These expressions can now be substituted in (1), and the two expansion coefficients written as

$$\alpha_x = (s_{11}+s_{12})q_x + s_{13}q_z, \qquad \text{......}(2a)$$

$$\alpha_z = 2s_{13}q_x + s_{33}q_z, \qquad \text{......}(2b)$$

where q_x and q_z are the so-called "thermal pressure coefficients" perpendicular and parallel to the hexagonal axis:

$$q_x = \frac{1}{v}\int_0^T \left(\frac{\partial C_p}{\partial \log l_x}\right)_T \frac{dT}{T}, \qquad \text{......}(3a)$$

$$q_z = \frac{1}{v}\int_0^T \left(\frac{\partial C_p}{\partial \log l_z}\right)_T \frac{dT}{T}. \qquad \text{......}(3b)$$

This represents the limit of purely thermodynamic reasoning. In order to evaluate q_x and q_z as functions of temperature, the quantum theory of specific heats must be used.

The heat capacity at constant volume, C_v, for a hexagonal crystal may be considered as being composed of two components, C_{v_x} and C_{v_z}, where

$$C_v = \tfrac{2}{3}C_{v_x} + \tfrac{1}{3}C_{v_z}. \qquad \text{......}(4)$$

These two components refer to vibrations perpendicular and parallel to the hexagonal axis, and are expressible by the Debye functions

$$C_{v_x} = 3RD\left(\frac{\Theta_x}{T}\right); \qquad C_{v_z} = 3RD\left(\frac{\Theta_z}{T}\right), \qquad \text{......}(5)$$

where Θ_x and Θ_z are the two characteristic temperatures referring to vibrations perpendicular and parallel to the hexagonal axis. Similarly, two components of the heat capacity at constant pressure may be defined by

$$C_p = \tfrac{2}{3}C_{p_x} + \tfrac{1}{3}C_{p_z}, \qquad \text{......}(6)$$

the two components again referring to the X and Z vibration directions. Hence, to a high degree of approximation, C_{px} is independent of small changes in l_z, and vice versa. We can therefore re-write the integrals (3) as

$$q_x = \frac{2}{3v}\int_0^T \left(\frac{\partial C_{px}}{\partial \log l_x}\right)_T \frac{dT}{T}, \qquad \text{......}(7a)$$

$$q_z = \frac{1}{3v}\int_0^T \left(\frac{\partial C_{pz}}{\partial \log l_z}\right)_T \frac{dT}{T}. \qquad \text{......}(7b)$$

In order to proceed further, C_{p_x} and C_{p_z} must be replaced by expressions involving C_{v_x} and C_{v_z} which, when written as Debye functions, are susceptible to manipulation.

The well-known identity

$$C_p = C_v + \frac{\alpha^2 vT}{\kappa}$$

may be approximated by

$$C_p = C_v + GvT,$$

where G is a constant independent of volume and temperature, since the temperature variations of α, the volume expansion coefficient, and of κ, the volume compressibility, are of the same order of magnitude over the higher temperature range. Hence

$$C_{p_x}=C_{v_x}+G_1vT\ ;\quad C_{p_z}=C_{v_z}+G_2vT,$$

where the constants G_1 and G_2 are defined by

$$G=\tfrac{2}{3}G_1+\tfrac{1}{3}G_2.$$

It is now easily shown that, since $v=l_x\,l_y\,l_z$,

$$\frac{\partial C_{p_x}}{\partial \log l_x}=\frac{\partial C_{v_x}}{\partial \log l_x}+G_1vT,$$

$$\frac{\partial C_{p_z}}{\partial \log l_z}=\frac{\partial C_{v_z}}{\partial \log l_z}+G_2vT.$$

The integrals in the expressions (7) for q_x and q_z therefore become

$$\int_0^T\left(\frac{\partial C_{v_x}}{\partial \log l_x}\right)_T\frac{dT}{T}+G_1vT \qquad \text{......(8 a)}$$

and

$$\int_0^T\left(\frac{\partial C_{v_z}}{\partial \log l_z}\right)_T\frac{dT}{T}+G_2vT. \qquad \text{......(8 b)}$$

To evaluate the C_{v_x} integral in (8 *a*), we express the heat capacity as the Debye function

$$C_{v_x}=\quad 3R\,D\left(\frac{\Theta_x}{T}\right),$$

whence

$$\begin{aligned}\frac{\partial C_{v_x}}{\partial \log l_x}&=\quad 3R\,D'\left(\frac{\Theta_x}{T}\right)\cdot\frac{1}{T}\cdot\frac{d\Theta_x}{d\log l_x}\\ &=\quad 3R\,D'\left(\frac{\Theta_x}{T}\right)\cdot\frac{\Theta_x}{T}\cdot\frac{d\log\Theta_x}{d\log l_x}\\ &=-\,3R\gamma_x D'\left(\frac{\Theta_x}{T}\right)\cdot\frac{\Theta_x}{T},\end{aligned} \qquad \text{......(9)}$$

where $D'\left(\frac{\Theta_x}{T}\right)$ is the derivative of the Debye function with respect to $\frac{\Theta_x}{T}$, and where

$$\gamma_x=-\frac{d\log\Theta_x}{d\log l_x}$$

may be considered as a numerical constant independent of temperature (Grüneisen number). It can now be shown from (9), following a similar procedure to that of Grüneisen and Goens, that

$$\begin{aligned}\int_0^T\left(\frac{\partial C_{v_x}}{\partial \log l_x}\right)_T\frac{dT}{T}&=3R\gamma_x D\left(\frac{\Theta_x}{T}\right)\\ &=\gamma_x C_{v_x}.\end{aligned} \qquad \text{......(10 a)}$$

Similarly,

$$\int_0^T\left(\frac{\partial C_{v_z}}{\partial \log l_z}\right)_T\frac{dT}{T}=\gamma_z C_{v_z}, \qquad \text{......(10 b)}$$

where

$$\gamma_z=-\frac{d\log\Theta_z}{d\log l_z}.$$

From the relationships (7), (8) and (10) we can now express the thermal pressure coefficients in a form suitable for evaluation :

$$q_x = \frac{2}{3}\frac{\gamma_x}{v}C_{v_x} + \frac{2}{3}G_1 T, \qquad \ldots\ldots(11\,a)$$

$$q_z = \frac{1}{3}\frac{\gamma_z}{v}C_{v_z} + \frac{1}{3}G_2 T. \qquad \ldots\ldots(11\,b)$$

At low temperatures the terms in T can be neglected and the relationships become simply

$$q_x = \frac{2}{3}\frac{\gamma_x}{v}C_{v_x}\,; \quad q_z = \frac{1}{3}\frac{\gamma_z}{v}C_{v_z}. \qquad \ldots\ldots(12)$$

The expressions derived by Grüneisen and Goens for low temperatures,

$$q_x = \frac{\overline{\gamma}_x}{v}3\boldsymbol{R}\,D\left(\frac{\overline{\Theta}_x}{T}\right); \quad q_z = \frac{\overline{\gamma}_z}{v}3\boldsymbol{R}\,D\left(\frac{\overline{\Theta}_z}{T}\right), \qquad \ldots\ldots(13)$$

are essentially similar to the approximate expressions (12). Their definition of $\overline{\gamma}_x$, $\overline{\gamma}_z$, $\overline{\Theta}_x$ and $\overline{\Theta}_z$ is, however, more complex than the simple definition of the similar quantities in this paper, and it is not certain that they have quite the same significance. A comparison of (12) and (13) shows that the γ's differ by factors of $\frac{2}{3}$ and $\frac{1}{3}$ if the Θ's are taken the same.

Substitution of the values (11 *a*), (11 *b*) for q_x and q_z in equations (2 *a*) and (2 *b*) gives the final form of the expressions for the two expansion coefficients $\alpha_\perp$ and $\alpha_\parallel$. These may conveniently be written as

$$\alpha_\perp = \alpha_x = AC_{v_x} + BC_{v_z} + CT, \qquad \ldots\ldots(14\,a)$$

$$\alpha_\parallel = \alpha_z = LC_{v_x} + MC_{v_z} + NT, \qquad \ldots\ldots(14\,b)$$

where

$$A = \frac{2}{3}\frac{\gamma_x}{v}(s_{11}+s_{12}), \qquad \ldots\ldots(15\,a)$$

$$B = \frac{1}{3}\frac{\gamma_z}{v}s_{13}, \qquad \ldots\ldots(15\,b)$$

$$L = \frac{4}{3}\frac{\gamma_x}{v}s_{13}, \qquad \ldots\ldots(15\,c)$$

$$M = \frac{1}{3}\frac{\gamma_z}{v}s_{33}, \qquad \ldots\ldots(15\,d)$$

$$C = \frac{1}{3}\{2G_1(s_{11}+s_{12}) + G_2 s_{13}\}, \qquad \ldots\ldots(15\,e)$$

$$N = \frac{1}{3}\{4G_1 s_{13} + G_2 s_{33}\}. \qquad \ldots\ldots(15\,f)$$

In computing, A, B, C, L, M and N can be treated as constants.

§3. THERMAL EXPANSION OF GRAPHITE

Magnus (1923) has measured C_p for graphite over the temperature range −229° to 827° C. and has derived the corresponding values for C_v. Using expressions equivalent to equations (4) and (5) of this paper, he found values of Θ_x and Θ_z which gave a good representation of C_v over the whole range. These are

$$\Theta_x = 2280\,; \quad \Theta_z = 760.$$

These values are used in the following calculations as they are in agreement with our definition and use of the quantities. The experimental data used here have been listed in Part I.

Thermal expansion of the a *dimension*

The experimental data are such that a reliable plot of $\alpha_\perp$ against temperature cannot be obtained. It is preferable, therefore, to consider the *a* unit-cell dimensions themselves.

The expression (14 *a*) for $\alpha_\perp$ will be approximated by

$$\alpha_\perp = AC_{v_x} + BC_{v_z}, \qquad \ldots\ldots(16)$$

which assumes that the term directly proportional to the temperature can be neglected. There is some justification for this as the constant C is composed of positive and negative components (s_{13} is negative), both of which are of the same order of magnitude. s_{13} is known to be negative because of the observed shrinkage in the *a* dimension (see the qualitative discussion in Part I).

Integration of (16) leads to the following expression in a:

$$\log a = AU_x + BU_z + \log a_0, \qquad \ldots\ldots(17)$$

where a_0 is the value of a at absolute zero, and where U_x and U_z, being the integrals of C_{v_x} and C_{v_z}, are components of the internal energy U. U_x and U_z can be evaluated by expressing them in terms of the Debye energy functions which have been tabulated by Beattie (1926); thus

$$U_x = 3RT\,F\left(\frac{\Theta_x}{T}\right); \quad U_z = 3RT\,F\left(\frac{\Theta_z}{T}\right),$$

where $F\left(\frac{\Theta}{T}\right) = \frac{U}{3RT}$ is the tabulated function.

The problem now is to choose values for A, B and a_0 which give the best agreement with the experimental data. The curve of a against temperature shown in figure 1 was calculated from (17), using the following values:

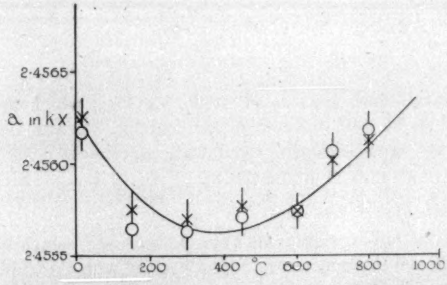


Figure 1. Comparison of observed values of *a* with the theoretical curve. OX, experimental values from two sets of measurements. Limits of experimental error indicated by vertical lines.

$$A = 0{\cdot}1620 \times 10^{-13}; \quad B = -0{\cdot}1013 \times 10^{-13}; \quad a_0 = 2{\cdot}45671 \text{ kx}.$$

The experimental values (two sets of measurements—see Part I), with their limits of error, are also indicated. The agreement is as good as can be expected considering the limitations of the experimental data.

From the above values for A and B, a plot of $\alpha_\perp$ against temperature can be derived from equation (16), using Beattie's tables of $D\left(\frac{\Theta}{T}\right)$ to evaluate C_{v_1} and C_{v_2}. The curve is shown in figure 2. It is in general agreement with experiment, particularly in giving a somewhat greater value for the negative coefficient at 0° c. than for the positive coefficient at 800° c.

Temperature (°c.)	$\alpha_\perp$ observed	$\alpha_\perp$ calculated
0	*ca* $-1{\cdot}5\times10^{-6}$	$-1{\cdot}3\times10^{-6}$
800	*ca* $0{\cdot}9\times10^{-6}$	$0{\cdot}8\times10^{-6}$

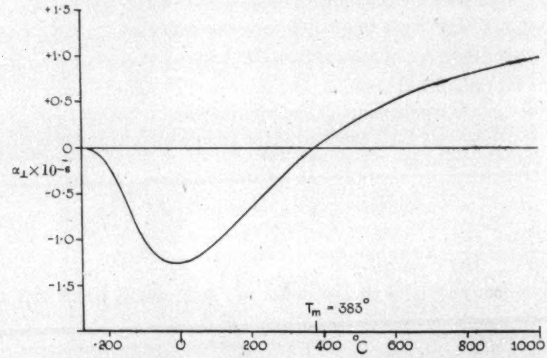


Figure 2. Theoretically derived curve for $\alpha_\perp$.

The theoretical curve gives $\alpha_\perp=0$ at $T_m=383$° c., which agrees with the temperature of about 400° c. found experimentally. It can be said that the theoretical curves agree reasonably well with experiment with regard to the variation with temperature of both a and $\alpha_\perp$.

Thermal expansion of the c *dimension*

In this case the full expression (14 *b*) for $\alpha_\parallel$ was used, which includes a term directly proportional to temperature. The best fit with the experimental data was obtained with the following values for L, M and N:

$$L=-0{\cdot}770\times10^{-13}\,;\quad M=1{\cdot}380\times10^{-13}\,;\quad N=1{\cdot}077\times10^{-8}.$$

The resulting theoretical curve is compared with the experimental values for $\alpha_\parallel$ in figure 3. The agreement is close. Of particular interest is the step-like character of the curve, which was seen in Part I to be demanded by the experimental data. Theoretically, $\alpha_\parallel$ must rapidly diminish at temperatures lower than about 400° K., and must, of course, be zero at 0° K.

As might be expected, it proved impossible to obtain agreement with experiment when the term NT in (14 *b*) was neglected. With such an expression, the

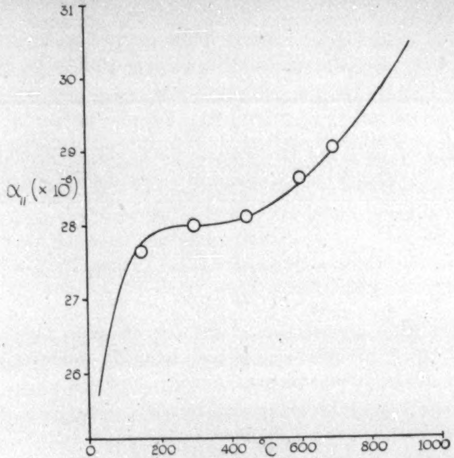


Figure 3. Comparison of experimental values for $\alpha_{\parallel}$ with the theoretical curve. ○, experimental values.

"best" curve for $\alpha_{\parallel}$ rose sharply to a maximum at about 500° K. and then decreased gradually with increasing temperature.

Thermal expansion at very high temperatures

At very high temperatures, the Debye functions approach the value unity. This is expressed more exactly by saying that

$$D\left(\frac{\Theta}{T}\right) \to 1 \quad \text{when } \frac{\Theta}{T} \to 0.$$

For graphite, this condition is reached approximately when $T > 2000°$ K. (1700° C.). Above this temperature

$$C_v \approx C_{v_x} \approx C_{v_z} \approx 3R.$$

Hence the expansion coefficient can be expressed approximately as a linear function of T:

$$\alpha_{\parallel} \approx (L+M).3R + NT$$
$$\approx 15 \times 10^{-6} + 1{\cdot}1 \times 10^{-8}T. \qquad \ldots\ldots(18)$$

Similarly

$$\alpha_{\perp} \approx (A+B).3R + CT$$
$$\approx 1{\cdot}5 \times 10^{-6} + CT. \qquad \ldots\ldots(19)$$

In this case, if C is negligible, $\alpha_{\perp}$ is approximately constant at very high temperatures, and has a value of about $1{\cdot}5 \times 10^{-6}$.

§4. ELASTIC MODULI AND COMPRESSIBILITIES

The values derived above for A, B, L and M can be used to derive an approximate estimate of the elastic moduli $(s_{11}+s_{12})$, s_{13} and s_{33}, if one additional datum relating these quantities is known. Such a datum is provided by the volume compressibility. In the hexagonal system

$$\kappa = 2\kappa_\perp + \kappa_\parallel,$$

where $\kappa_\perp$ and $\kappa_\parallel$ are the linear compressibilities perpendicular and parallel to the hexagonal axis. They are related to the elastic moduli by the expressions

$$\kappa_\perp = s_{11} + s_{12} + s_{13}\ ; \quad \kappa_\parallel = 2s_{13} + s_{33},$$

whence

$$\kappa = 2(s_{11} + s_{12}) + 4s_{13} + s_{33}.$$

In addition, from A, B, L and M we obtain the following quantities by using relations (15 *a*) to (15 *d*) with $v = 5{\cdot}3$ (volume of a gram-atom of graphite at 18° c.):

$$\begin{aligned} \gamma_x(s_{11} + s_{12}) &= 1{\cdot}29 \times 10^{-13}\ \text{cm.}^2/\text{dyne}, \\ \gamma_z s_{13} &= -1{\cdot}61 \times 10^{-13}\ \text{cm.}^2/\text{dyne}, \\ \gamma_x s_{13} &= -3{\cdot}06 \times 10^{-13}\ \text{cm.}^2/\text{dyne}, \\ \gamma_z s_{33} &= 21{\cdot}94 \times 10^{-13}\ \text{cm.}^2/\text{dyne}. \end{aligned}$$

Using Basset's (1941) measurement of the instantaneous compressibility at 18° c., $\kappa = 44{\cdot}9 \times 10^{-13}$ cm.^{2}/dyne, we can now derive the following approximate values for the elastic moduli and the two Grüneisen numbers :

$$\begin{aligned} s_{11} + s_{12} &= 1{\cdot}8 \times 10^{-13}\ \text{cm.}^2/\text{dyne at } 18^\circ\ \text{c.}, \\ s_{13} &= -4{\cdot}3 \times 10^{-13}\ \text{cm.}^2/\text{dyne at } 18^\circ\ \text{c.}, \\ s_{33} &= 58{\cdot}5 \times 10^{-13}\ \text{cm.}^2/\text{dyne at } 18^\circ\ \text{c.} \end{aligned}$$

$$\gamma_x = 0{\cdot}71\ ; \quad \gamma_z = 0{\cdot}38.$$

The two linear compressibilities are, approximately,

$$\kappa_\perp = -2{\cdot}5 \times 10^{-13}\ ; \quad \kappa_\parallel = 50 \times 10^{-13}\ \text{cm.}^2/\text{dyne}.$$

There is thus a small negative compressibility perpendicular to the hexagonal axis. The anisotropy of the compressibility is seen to be considerable, the effect of pressure being mainly confined to packing the graphite layers closer together. Values of s_{11} and s_{12} separately are, unfortunately, not obtainable from our data. It is, however, of some interest to have for such an intractable material as graphite even approximate estimates of certain of the elastic moduli.

§ 5. COMPARISON WITH ZINC AND CADMIUM

It is of interest to compare the data for graphite with those obtained for zinc and cadmium by Grüneisen and Goens. Table 1 gives the values of the elastic moduli, compressibilities and thermal expansion coefficients which refer to a temperature of about 18° c.

Table 1

	$s_{11}+s_{12}$	s_{13}	s_{33}	$\kappa_\perp$	$\kappa_\parallel$	$\alpha_\perp$	$\alpha_\parallel$
	($\times 10^{-13}$ cm.^{2}/dyne)			($\times 10^{-13}$ cm.^{2}/dyne)		($\times 10^{-6}$)	
Zinc	7·5	−6·1	28·2	1·4	16·0	13	64
Cadmium	10·8	−9·3	35·5	1·5	16·9	11	100
Graphite	1·8	−4·3	58·5	−2·5	50	−1·2	26

The interesting point emerges that whereas $\kappa_{\parallel}$ is much greater for graphite than for zinc or cadmium, $\alpha_{\parallel}$ is much smaller. This is contrary to the usual state of affairs, as is seen, for instance, by comparing the data for zinc and cadmium. An explanation along the following lines is suggested. Graphite is an extreme example of a layer structure, so much so that it is an assembly of large macromolecules held together by what are probably mainly van der Waals forces. The covalent linking within a layer is so strong that the layer can be considered to be almost rigid. Thermal vibrations, perpendicular to the layers, then, will be diminished in amplitude owing to the large mass of each layer, and the thermal expansion will not be as great as the extreme anisotropy of the structure would at first suggest. The effect of pressure, however, is independent of the mass of the layer, and a high compressibility results because of the large interlayer spacing (3·35 A.) and weak interlaying bonding.

The great strength of the C—C bonds within a layer is reflected in the small value of $s_{11}+s_{12}$, whereas the weak interlayer forces give rise to the large value for s_{33}.

The partially covalent nature of the bonds between atoms in the same layer in zinc and cadmium is similarly demonstrated.

Another interesting comparison involves the ratio of interatomic distances within and between layers. In table 2, d_1 is the interatomic distance in A. within a layer, and d_2 the closest distance of approach of atoms in neighbouring layers.

Table 2

	d_1	d_2	d_2/d_1	$\kappa_{\parallel}/\kappa_{\perp}$
Zinc	2·66	2·91	1·09	11·4
Cadmium	2·97	3·29	1·11	12·1
Graphite	1·42	3·35	2·36	−20

The ratio of the linear compressibilities is obviously correlated with the ratio $\frac{d_2}{d_1}$, while the negative sign of $\frac{\kappa_{\parallel}}{\kappa_{\perp}}$ for graphite is one more indication of the great anisotropy of the structure.

§ 6. ACKNOWLEDGMENTS

The work constitutes part of a research programme of the British Coal Utilisation Research Association carried out under the direction of Dr. D. H. Bangham, who has been unfailing in his help and encouragement. The writer is also indebted to Dr. A. J. C. Wilson for his generous assistance on several points, and to Mr. A. H. Wilson for his critical reading of the manuscript.

REFERENCES

Basset, J., 1941. *C.R. Acad. Sci., Paris*, **213**, 829.
Beattie, J. A., 1926. *J. Math. Phys.* **6**, 1.
Grüneisen, E. and Goens, E., 1924. *Z. Phys.* **29**, 141.
Magnus, A., 1923. *Ann. Phys., Lpz.*, **70**, 303.
Voigt, W., 1910. *Lehrbuch der Kristallphysik* (Leipzig and Berlin: Teubner).

STRUCTURE AND ORIENTATION IN THIN FILMS OF POLYTHENE

By A. CHARLESBY,
Imperial College of Science and Technology, London

MS. received 3 *April* 1945

§ 1. INTRODUCTION

An extended series of investigations into the crystal structure and habits of a number of long-chain saturated hydrocarbons has been published, generally limited to paraffins or their derivatives containing from 10 to 50 carbon atoms. While the data obtained in this range constitute an invaluable basis on which to found theoretical studies of intermolecular binding forces, some difficulty arises due to the presence of end groups. A considerable simplification results when the length of each molecule may be considered as infinite, so that each atom bears the same relationship to its neighbours. In common with a number of other properties, the structure of such a hydrocarbon can in a certain measure serve as a limit towards which existing structures approach with increasing length of the CH_2 . . . chain.

Bunn (1939) has made use of x rays to investigate hydrocarbons of this character and has deduced the electron density of the structure. In the present work, thin films of similar high polymers were studied by means of electron diffraction. In general, good agreement is found in those aspects of the problem which overlap in the two methods of investigation.

A highly polymerized hydrocarbon, polythene, was available for these experiments. It is produced by the polymerization of ethylene, and occurs in a range of molecular weights, which, according to viscosity measurements, amount to several thousand carbon atoms. The commercial product may be obtained in several forms with different average degrees of polymerization. As will be shown later, no appreciable difference was detected between films formed from these different specimens. No attempt was made to obtain specimens whose molecules cover only a very narrow range of molecular weights.

At room temperature the polymer is a white solid, reminiscent of paraffin wax in appearance, but its high degree of polymerisation results in a very tough structure. It shows no definite melting point but softens in the temperature range 90–100° c. At slightly lower temperatures it loses its wax-like appearance and becomes more transparent. As Garner, Bibber and King (1931) have shown, a similar change may be observed in the transition of docosane from the β to α form, which is stable at temperatures several degrees below the melting point. A general formula was deduced by these investigations for the setting point of a hydrocarbon C_nH_{2n+2}:—

$$T=(0{\cdot}6085n-1{\cdot}75)/(0{\cdot}001491n+0{\cdot}00404).$$

For a molecule of infinite length $T = 408°$ K., a temperature somewhat higher than the softening range observed in polythene.

In the present paper, the structure and orientation of polythene at room temperature are discussed. In the following paper the influence of temperature on the structure and the changes which occur in the neighbourhood of the melting point are described.

§2. ELECTRON DIFFRACTION—EXAMINATION OF THIN POLYTHENE FILMS

Preparation of specimens

For the purpose of these experiments, polythene was available in two solid forms known as A and B, which differ in their viscosity in solution, and hence in their mean molecular weight. In addition, experiments were carried out on polythene in the form of an emulsion ; these will be described later.

In view of the range in molecular weights, and the possible inclusion of grease or other impurities, some degree of purification was judged advisable. Solutions of polythene A (available in the form of rods) were prepared in hot xylene, from which the polythene was precipitated on cooling, in the form of a thick translucent gel. This process was repeated on a number of occasions and specimens from the successive filtrates examined by electron diffraction. In no case was there any appreciable change in either spacing or intensity of the patterns thus obtained, nor did these differ from patterns obtained from specimens of the precipitate. It may therefore be concluded that the polythene under examination does not contain very appreciable quantities of impurity and that it may be considered as a compound sufficiently homogeneous for structure determinations. The inclusion within a pure substance of 1% of a homologous impurity would not, however, have any great influence on the structure.

At first a solution of polythene in benzene was used, but a 0·5% solution of a precipitate from xylene proved far more satisfactory for the preparation of thin films, and this was generally adopted. One or two drops of the solution were placed on hot water (90°c.) and the xylene evaporated off. Films produced in this fashion measured several square centimetres in area, and their thickness, which was variable as judged by their colour, was of the order of 1000 A. over the major part of the area. The thinnest films produced were of a steely white or golden yellow colour, while thicker films presented a wide and pleasing range of colouring. The films found most useful for transmission were obtained from films of about 1000 A. thickness. These films are very coherent in spite of their thinness. This may readily be seen when tension is applied to one portion of the film, since the whole film tends to move as a single skin over the surface of the liquid.

When a suitable portion of the film was found, it was removed on a nickel gauze and mounted in the electron-diffraction camera. This was of the Finch type, making use of a cold-cathode discharge tube operating at about 50-60 KV. The distance from specimen to photographic plate is about 50 cm. To produce satisfactory blackening of the plate, exposures of 10–30 sec. were most suitable.

Elucidation of structure

Examples of the electron-diffraction patterns obtained with the film normal to the incident beam are shown in figure 1. The observed spacings are compared with those given by Bunn for a highly polymerized saturated hydrocarbon, as obtained from x-ray examination. The calculated spacings are based on the following axial dimensions :—

Bunn	$a=7\cdot40$,	$b=4\cdot93$,	$c=2\cdot534$.
Charlesby	$a=7\cdot428$,	$b=4\cdot934$,	$c=2\cdot532$.

In examination of these films the three axes may be taken as mutually perpendicular.

One pattern was obtained from a composite film of graphite and polythene, and this was used to standardize the more intense rings of the polythene pattern in terms of the 1·2217 A. graphite spacing. A number of other patterns were also measured in terms of the 2·467 polythene spacing, which was generally used as sub-standard. The difference from the mean spacings rarely amounted to more than 0·4% and was generally about 0·2%.

Unlike x-ray diffraction patterns of polythene, electron diffraction patterns show sharp rings, there being little or no trace of haloes due to amorphous material. This difference is to be ascribed to the different types of specimens used; films of about 1000 A. thickness prepared as above are essentially crystalline, whereas polythene prepared from the melt in thick specimens and in the absence of an orientating substratum consists to a considerable extent of amorphous material. The extent of agreement shown for the structure of molecules which may be considered infinite in extent may be seen by comparing them with the corresponding spacings typical of shorter C_nH_{2n+2} chains.

Table 1

n in C_nH_{2n+2}	a	b	
19	7·55	5·01	Müller (1932)
23	7·43	4·97	,,
30	7·33	4·92	,,
31	7·40	4·93	,,
34	7·40	4·95	,,
44	7·33	4·93	,,
35	7·43	4·97	Hengstenberg
30	7·51	4·955	Schoon
31	7·49	4·97	,,

In these measurements a change of 5° c. would cause an error of the order of 1% in the spacings of a and b.

This comparison of the spacings of polythene with those of saturated hydrocarbons shows that the lateral packing of the molecules must be very similar in the two cases. It immediately follows that the polymerized molecule is linear in character.

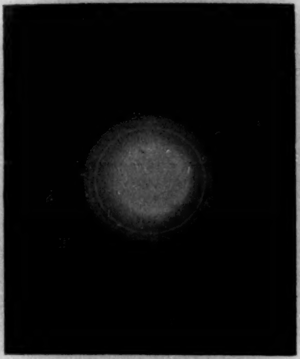


Figure 1.
Polythene film normal to beam.

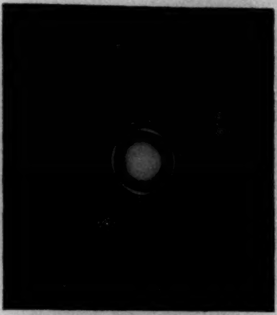


Figure 2.
Polythene film inclined to incident beam.

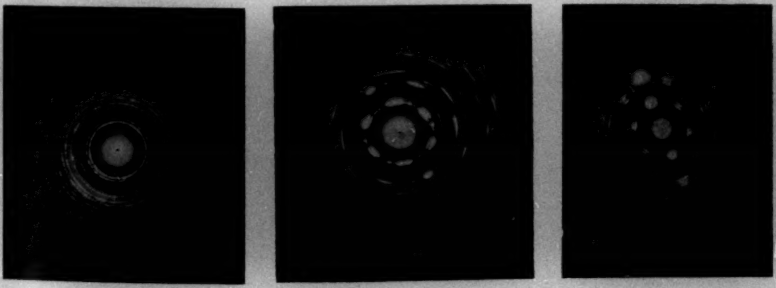


Figure 3. Paraffin wax films grown on water at temperatures of 20° c., 52° c., 100° c.

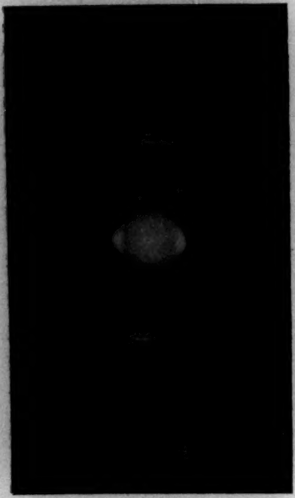


Figure 4.
Polythene film fully extended.

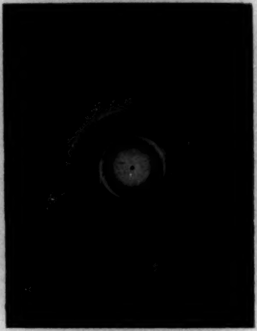


Figure 5.
Polythene film partially extended.

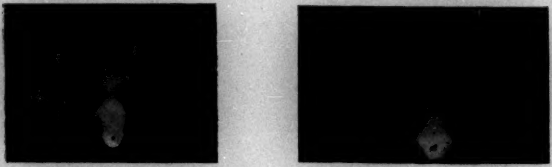


Figure 6. Orientation of polythene on a steel surface.
(Incident beam perpendicular and parallel to direction of application.)

This similarity in structure, taken in conjunction with the sharpness of the rings (arising from the relatively large crystal size) and the absence of haloes due to amorphous material, provides strong evidence that during and after polymerization the vast majority of the polythene molecules retain a linear structure without side-chains. The intrusion of a single methyl side-group would of course disrupt the crystal over a wide extent.

While the spacings observed are essentially similar to those of the saturated hydrocarbon studied by Bunn and others, the intensity distribution as between the rings shows considerable differences. For example, among the shorter spacings, the intensity of the 200 is observed to be far weaker and more diffuse than that of the 110 and the 020 diffractions.

Orientation

The above results were deduced with the incident electron beam normal to the polythene film. To investigate the orientation which accounts for the modified intensity distribution, specimens were tilted through angles of between 20° and 70°. The pattern then showed considerable differences in directions parallel to, and at right angles to, the axis of orientation. These are given in table 2 for certain diffractions of the lowest order. It will be seen that while 020 is weakened if the film is rotated, 200 is intensified. Thus in the film as

Table 2. Intensities of rings (visual estimates)

Plates 4006, 4032 : Film normal to incident beam.
Plates 4008, 4034 : Film inclined at about 30°.
Intensities in directions parallel and at right angles to axis of rotation.

Plane	4006	4032	4008		4034	
110	VS	VS	VS	S*	VS	VS
200	M	M	M	S	MW	S
210	F	—	—	—	VF	VF
020	S	S	S	—	S	M
120	F	F	—	—	F	—
011	MS	MS	M	M	M	MW
310 110	M	MW	MW	MS	—	—
201 220	M	MS	M	M	MS	M
211	MW	MW	MW	MW	W	MW
320 410 121	MS	MS	MS	MW	M	M
311	W	W	—	—	W	M

* 110 ring in 4008 (which shows orientation most clearly) is split into four very intense quadrants, at about 45° to the axis of rotation.

The visual estimates of intensity are in the following order :—VS, S, MS, M, MW, W, VW, F, VF.

picked up from the water surface, the 020 plane is approximately perpendicular to the surface, but the 200 is inclined, i.e. the *b* axis lies parallel to the water surface while the *a* axis is inclined at some angle. This conclusion is verified by an examination of plate 4008, obtained from a specimen the perfection of whose orientation is well above average. Here the 110 diffraction is intensified in directions making angles of about 35° with the axis of rotation. The considerable length of the arcs shows that even in this specimen the orientation is by no means perfect.

Angle of tilt of a *axis*

Among the large series of paraffins, aliphatic acids and alcohols, the structures of which have been determined, several enantrotropic forms are known to occur. Müller's work on the hydrocarbons has shown that, in a number of cases, transition points occur at certain temperatures, corresponding to a change in structure. Francis, Collins and Piper moreover have noted that in a number of hydrocarbons, the crystallized form depends on the manner in which the crystals were obtained, e.g. on whether crystals were grown from a polar or non-polar solvent. These and similar observations are to be explained as due to the great similarity in the main structure of these hydrocarbons. The main binding force arises from the van der Waals and repulsion forces between neighbouring CH_2 groups. Because of their relative scarcity, side and end groups can only produce considerable modifications in structure under special conditions, e.g. due to steric hindrance. Where this perturbation is not serious, the main chains pack parallel to one another at almost constant distances with the CH_2 groups interleaving. The structure thus obtained is very stable to longitudinal shearing stress, except when its magnitude is such that the CH_2 groups in neighbouring chains ride over one another by an integral number of zig-zags, in which case the binding energy is only slightly altered. In principle, one may therefore anticipate a series of discrete stable or meta-stable structures in each of which the spacings at right angles to the chain length are unaltered.

§ 3. MOLECULAR ARRANGEMENT OF POLYTHENE

The anomalous intensity distribution observed in the orientation of polythene may be readily explained in terms of the above considerations, if the structure of the polymerized molecule is considered to be substantially linear in character, but that in the course of polymerization two dissimilar end-groups are formed :—

$$CH_2=CH_2+CH_2=CH_2+\ \ldots\ \rightarrow CH_3-CH_2-CH_2-\ \ldots -CH=CH_2.$$

While the absence of side-chains permits a lateral packing similar to that prevailing in the saturated hydrocarbons, the unsaturated end-group is too large to be accommodated within the main body of the structure.

In a dilute solution on a water surface the individual molecules orient themselves with the hydrophilic group pointing down into the water. As the solvent evaporates the molecules of the polymer tend to crystallize in the manner customary for a saturated hydrocarbon, but are prevented from so doing by the steric hindrance of the end-groups. This hindrance is, however, avoided in certain monoclinic and triclinic modifications and the molecules are consequently tilted towards the water surface.

In stearic acid films grown on water, two forms α, β are known to occur with the molecules inclined at definite angles to the normal, depending on whether *b* or *a* axis remains on the water interface. In the large number of polythene specimens examined, only the α form (our notation {*n*O}) has been observed.*

While the end-group may assume a number of different positions, it must move as a rigid unit if there is to be no distortion of bond spacings or angles. The only variation which appears feasible is one of rotation about the penultimate carbon-carbon bond.

Three special cases may be considered :—

(*a*) All carbon atoms lie in the main plane of the molecule, and approximately along its axis.

(*b*) All carbon atoms lie in the main plane of the molecule, but the C=C bond is approximately at right angles to the axis of the molecule.

(*c*) All the atoms of the $-CH{=}CH_2$ group lie in a plane at right angles to the axis of the molecule.

In the saturated hydrocarbon, the closest distance of approach between atoms in different molecules occurs between two hydrogen atoms. The repulsion arising from the interaction between these two atoms plays an outstanding rôle in stabilizing the structure.

In all three cases outlined above, the minimum distance of approach between atoms in adjacent molecules is very much smaller than the value for the saturated hydrocarbon, which is about 2·5 A. (assuming Bunn's data and a CH bond-length of 1·09 A.). Even in the most favourable case (*a*) this distance is only slightly greater than 2·0 A. To avoid the excessive repulsion arising from this close proximity, while yet retaining the stability inherent in the main hydrocarbon-like structure, adjacent molecules may be sheared to avoid the steric interference. The monoclinic structure which successfully avoids these excessive repulsive forces is the 20. In this structure the angle between the base and the *c* axis is 56°.

There is no steric hindrance in the *b* axis direction, and no modification is therefore required. This result is in agreement with the observed structure in which only the angle between the *a* axis and the water interface is altered from 0°.

That the orientation in polythene is due to the influence of the end-group is rendered even more probable by an examination of the electron pattern obtained from paraffin films produced under similar conditions. While the spacings are similar to those of polythene, the intensity distribution shows very significant differences, due to the {0,0} structure of paraffin wax.

Structure of the film

The structure of the film is defined in so far as the side packing of molecules and the orientation at the base are concerned, but the variation in chain length of the individual molecules precludes the measurement of a crystal height. It appears highly probable that odd fragments or portions of polythene chain protruding from the main structure of each crystal link up with other crystals, forming a dense complex of crystals matted together by molecules common to two or more of them. The general orientation of the chains produced by the basic layer is thus to a considerable extent conserved. Nevertheless, some departure from the strict orientation must be anticipated, departure which becomes more serious in thicker specimens as the influence of the well oriented base layer becomes less effective because more remote. In the thicker films as prepared for X-ray investigation this influence is almost entirely lost in the mass of the solid. For thin films as used in electron diffraction, where the thickness generally corresponds to less than 1000 C—C linkages, the orientation is retained almost in its entirety. Even here, however, the orientation is to some extent impaired as is made clear by the appearance of orders which would be absent in a perfectly oriented crystal.

---

* Further evidence for the existence of an unsaturated end-group is provided by the discoloration which polythene produces in dilute $KMnO_4$.

A point which still remains obscure is related to the length of the individual molecules. In the A polymer the molecule is believed to comprise far more than 1000 C—C groups. It is then difficult to account for the observed structure, without assuming that the molecules are bent back on themselves, and this would presumably disorganize the crystal orientation very considerably. The observed orientation of the film, while by no means perfect, is sufficiently marked to make this hypothesis unlikely.

Influence of the rate of crystallization

Of considerable importance to the structure of the film is the rate and temperature at which crystallization takes place. The films of polythene used in these experiments were prepared on a water surface at 95° C. and formed a skin several square centimetres in size. At lower surface temperatures, when the mobility of the molecules on the water surface is reduced, the film formed is far more brittle, and readily breaks into a number of fragments. Films prepared under these conditions are far less readily mounted and examined.

An analogous effect may be observed in the case of the lower hydrocarbons. A thin film of paraffin wax, prepared from a drop of xylene solution on a water surface at 100° C., is found to consist of relatively large single crystals, measuring about 10^{-2} cm. across, all in approximately the same orientation. On a water surface at 52° C. the crystals formed are several times smaller, while at 20° C. they only measure 10^{-3} to 10^{-4} cm. across. The films of polythene examined all consisted of crystals estimated to measure about 10^{-3} to 10^{-4} cm. across.

Attempts were made to grow polythene crystals by evaporation of the solution on a glass surface. They all resulted in a conglomeration of small milky patches without mutual coherence. Better results might possibly be achieved on a warm glass surface.

Pseudo-hexagonal crystals of paraffin have been observed by crystallization near the melting point, instead of the usual orthorhombic variety, but no such crystals have as yet been observed in polythene films, nor does their existence appear likely.

Samples of B polythene (of lower average molecular weight than A polythene) appear to have a somewhat different texture and appearance. Specimens of B polythene film prepared as above were found to possess the same structure and orientation as A polythene, any slight differences in the patterns, e.g. relative intensities and width of rings, being adequately explained in terms of individual variations due to slight changes in the method of preparation and film thickness.

§4. DEFORMATION OF POLYTHENE FILM

When a plastic is deformed, and the stress is subsequently released, there is in general an instantaneous partial recovery, followed by a further slower recovery (after-effect). The final state of the specimen depends on its previous history, the rate and degree of extension, and the conditions under which the applied stress is released. In the study of the elasticity of polymers, temperature plays an exceptionally important rôle.

Polythene exhibits the property of cold-drawing to a remarkable degree. With the electron-diffraction camera it is possible to study the structural changes which take place during the cold-drawing of a thin film of polythene. Owing to the experimental difficulties, the degree of extension cannot be known accurately, but it is nevertheless possible to survey its influence on structure.

Thin films prepared as described above were picked up on a small jig consisting of two razor blades mounted in slides in which they could be moved. The initial separation of the knife edges was 0·3 mm. and this was increased during the extension to nearly 1 mm. During this extension part of the specimen was distorted or torn, so that the true extension could not be obtained. A pattern obtained from an extended film is shown in figure 4. In this case the initial colour of the film was yellow and white, the final colour being white.

In the initially unstressed film the molecules make angles of perhaps 60° to 120° with the direction of the subsequently applied force. Electron-diffraction patterns of the extended film show that the molecules assume an orientation nearly parallel to the applied stress, whereas the a and b axes are oriented in all directions at right angles to it. The spacings of the extended and unstretched material were found to be identical within the limits of experimental error.

As shown in figure 4 each hkl reflection consists of an arc. Details of the intensity distribution within each arc are more fully discussed below; it may be stated, however, that the length of each intense arc may be ascribed to imperfect orientation, so that the c axis of each crystal makes on the average an angle θ with the direction of the imposed stress. In one case this angle has been estimated from the length of the arcs. θ is calculated from the ratio $\tan\theta = l/d$, where l is the subtense of the arc, and d is the corresponding diameter. The results are shown below.

Table 3

Plane	004	112	002	121	110	200	020
l (mm.)	12	5	6	3·5	2	1·5	3
d (mm.)	72·4	37·6	36·4	26·7	11·2	12·7	18·7
$\tan\theta$	0·166	0·133	0·166	0·13	0·18	0·12	0·16

Mean value $= 9^\circ{\cdot}1 \pm 1^\circ$

Without a fuller investigation of intensities, it is not possible to state whether the *c* axes of different crystals all make angles of about 9° with the applied stress, or whether, as appears most likely, the angles vary from 0° to 90° according to some form of Maxwellian distribution.

§ 5. CHANGES IN THE STRUCTURE OF THE FILM DURING EXTENSION

The process involved in the extension of thin films of polythene may be deduced from a consideration of several patterns of films extended to varying degrees. In the intermediate stage it is found that 110, 200, 020, etc., diffractions become very weak in directions along the direction of extension, while 002 and corresponding diffractions become intensified. In the perpendicular direction 110, 200 and 020 retain their initial intensities almost unaltered. In intermediate directions the intensities vary in a continuous fashion from 0° to 90°. Thus the re-orientation of each micelle or crystallite takes place smoothly and not by a sudden change to substantially complete orientation.

In films which are almost completely extended, the 100, 200 and 020 are confined to short arcs, but within these the intensities show characteristic changes, which appear to rise from the superposition of short and very intense arcs on longer and weaker arcs. The more intense arcs subtend angles of about 8° ($\pm 4^\circ$ from the mean) with a relatively sharp end, while the weaker arcs subtend about 20°. In the 200 diffraction the weaker arc is extremely faint or missing.

The explanation of these observations is believed to be as follows :—The weaker and more extensive arcs correspond to remnants of the initial rings in the unstretched specimen, and are due to the relatively few micelles or crystals which during the extension remained almost unaltered in position. The stronger arcs arise from micelles whose position has been completely altered, and which finally come to rest in directions almost parallel to that of the extension.

As a result of this motion most micelles lie almost parallel to one another. It is believed that some form of partial recrystallisation takes place in this new orientation. This crystallization cannot be other than very imperfect, in view of the inherent strains and distortions to which the micelles are subject in their new positioning. Nevertheless the presence of very short arcs, which may almost be described as blobs or diffuse spots, and whose ends are relatively sharply defined, does lead one to envisage the existence of intermicellar forces which operate mainly between micelles brought into approximately parallel orientation by the applied stress, and which tend to complete the perfection of orientation when the external stress becomes almost inoperative. It appears very plausible to suppose that the completed blocks, however imperfectly crystallized, are fairly stable, so that the extended film does not revert to its original structure. The intermicellar forces are van der Waals forces which also serve to bind molecules within each micelle. The resultant force must, however, be weaker than the forces within each crystal, due both to the intermolecular strains introduced during the extension, and the greater intermolecular distances generally involved.

When the temperature is raised to a point slightly below the melting point of the pure crystal, these forces no longer suffice to maintain the rigidity of the structure, and there is a resultant tendency for the film to revert at least partially to the initial state.

The reason for the almost complete disappearance of the longer 200 arc is not very clear. A possible explanation is that this may be due to the initial orientation present in the film. The crystals are generally tilted so that the *a* axis is not parallel to the surface of the film. When the stress is applied, the first tendency is for this tilt to increase. Thus 200 reflections decrease in intensity more rapidly than those due to 020.

In the hypothesis outlined above, it is assumed that each crystal initially present in the film remains substantially unaltered during the extension. Other types of deformation may also be considered. The possibility of a shear within each crystal is rendered unlikely by the observed intensification, as well as by other considerations. An alternative hypothesis, that the extension consists mainly of a rearrangement of amorphous structure within the film is, at least for the thin films studied here, definitely disproved.

B polythene

Extended films of the lower polymer have been obtained but no marked difference in their structure could be detected. From the few specimens prepared, it appears that extended films of B polythene are somewhat less readily obtainable.

§ 6. ORIENTATION OF POLYTHENE

Orientation on metals

The experiments described above are limited to the orientation of polythene on water surfaces and, therefore, provide little evidence for possible differences in orientation produced under different surface conditions. A short series of experiments was therefore undertaken to throw some light on the influence of other substrata and of the treatment to which the film was subjected. Use was made of a metallic base, which offers advantages both from the practical interest in the adhesion of plastic to metals, and for the ease with which several forms of treatment may be applied.

Several experiments were made on a polished copper disc, but in most cases a chromium-plated steel disc was used. The experiments all tend to illustrate the lack of powerful orienting force at the surface, although adhesion is always sufficient to retain some polythene, even when the specimen is subjected to strong rubbing. The results and conclusions are summarized below :—

(*a*) A chromium-plated steel disc, previously polished, was cleaned by continued washing in benzene. Polythene was then rubbed on the disc in one direction, use being made of a suitably pointed stick of the polymer. The specimen was examined with the electron-diffraction beam parallel to and perpendicular to the direction of rubbing. The two patterns obtained are essentially different in character. With the incident beam parallel, the pattern consists of several rather diffuse rings, whereas with the beam perpendicular it comprises a number of *l* layer lines, with arcs such as 110, 200, 020, 011, 002, and 022 superimposed. The film is thus found to consist of a series of very small or imperfect crystallites, in which the preponderant orientation is imposed by the molecules which tend to lie along the direction of rubbing, as might indeed be anticipated on purely mechanical lines. The degree of one-dimensional orientation thus produced is fairly considerable, so that a number of *hk*0, *hk*1, *hk*2 reflections can be clearly distinguished. From the lengths of the arcs involved, it appears that the crystallites make angles of about 10° with the surface (or, alternatively, that the surface is not plane to this extent, although this is far less likely). In addition a very appreciable part of the film consists of less well-oriented crystallites, in which the *c* axes lie in all directions parallel to the surface.

To within experimental error the spacings involved are identical with those found in films produced by slow evaporation on the water surface, where conditions are far more favourable to crystallisation. On the other hand, the crystals formed are far less perfect or are considerably smaller.

(*b*) The specimen examined in (*a*) was subjected to rubbing on cotton wool in the same direction as the polythene had previously been applied. Examination showed that the orientation first found had become even more marked. In particular the rings produced by crystallites in random orientation (but with their *c* axes parallel to the surface) had disappeared ; thus the vast majority of the molecules in the film form crystallites in which the molecular chains lie in directions nearly parallel to each other.

(*c*) The specimen was then rubbed lightly with cotton wool in a direction perpendicular to the initial direction. The intensity of the pattern in (*b*) is considerably reduced ; in addition intense horizontal layer lines are observed. These are ascribable to a re-orientation of some of the polythene, with the chain standing erect on the substrate. This may either be due to liquefaction of certain regions, with subsequent recrystallisation in a more stable orientation, or possibly to a mechanical re-orientation of the crystals without melting. This latter hypothesis would appear most unlikely.

(*d*) A further surface as in (*a*) was prepared and subsequently subjected to hard rubbing with cotton wool in a perpendicular direction. From examination of patterns taken with the electron beam parallel to, and at right angles to, the initial direction of rubbing, it appears that this final application of friction has succeeded in rotating the crystals through $\pi/2$ into the final direction. From the diffuse character of the oriented chains, one may assume the mean length of the crystal (assumed perfect) to be of the order of 100 A.

These observations can be readily explained in terms of the weak adhesion between polythene film produced in the manner described and the chromium surface. At the same time the re-orientation produced leads one to favour the hypothesis that, by hard rubbing, liquefaction of polythene is produced with subsequently faulty recrystallization in the very limited time available before the temperature falls below a critical value. That this time is very short is shown by the fact that the molecules retain a very clear "memory" of the direction of rubbing, and that this determines their subsequent orientation during the process of crystallization.

(*e*) A further set of experiments investigated the structure of films produced by evaporation on water, and subsequently picked up on a steel surface. If a surface film of this nature is intensively dried, and rubbed with wool, most of the film is removed, leaving an invisible layer in which the crystallites are oriented with the c axes in the direction of the applied stress.

(*f*) After heating the thin film of polythene obtained in (*e*) to 180°C ($>$ melting point), the orientation is to some extent destroyed ; yet sufficient is retained to indicate quite clearly the initial direction of rubbing. When the film is further heated to 300° C., no trace of film can be detected.

Orientation of polished copper

Polythene was rubbed on the polished copper disc, and the specimens examined in directions both parallel and perpendicular to the direction of rubbing. The film produced is found to consist mainly of very small crystallites (estimated at about 10 molecules across) oriented in all directions, but with a slight preponderance of molecules standing normal to the surface. The film obtained was generally thin, and this may be related to the poor orientation observed. It must be stressed that the patterns are not as clear as those obtained from the steel specimen, nor is their interpretation so satisfactory. It would, however, appear that adhesion to polished copper is inferior to that of polythene on plated steel.

Orientation on acidified or alkaline water

In view of the orientation always found to occur on water surfaces, orientation which affects the whole crystal structure of the film, the question arises whether this may be changed by altering the pH of the water.

Films grown on hot acidified or alkaline water (HCl or NH_4OH) do show some weak extra rings, probably due to impurities. The main pattern consists of the usual rings, the spacings and intensities of which show no significant difference from those obtained in the main series of experiments.

§7. POLYTHENE ON COLLODION

A short experiment similar to that described above for metals was carried out on a collodion surface. A thin film of collodion prepared by evaporation of a drop of acetone solution on water was picked up on a rocksalt surface, which had previously been polished. On this collodion surface a film of polythene was prepared by evaporation of a drop of xylene solution. The rocksalt was then dissolved away and the composite film mounted on the gauze for examination.

The pattern of the film thus obtained consists of the usual collodion haloes and of the rings of polythene. The intensity distribution of the latter shows that the orientation is similar to that of polythene prepared on a water surface.

In another specimen the polythene film was rubbed prior to the removal of the rocksalt base. Contrary to expectation, no re-orientation of the film was observed. It must, therefore, be concluded either that the use of a collodion substratum prevents the application of considerable pressure, or that gross irregularities in the rocksalt surface protect most of the film.

In a third specimen, solid polythene was directly rubbed on to the collodion backing before the rocksalt base was dissolved away. The pattern of the film was found to correspond to a strongly oriented film in which the molecules lie in the direction of rubbing, and in which orders up to 004 may be detected. The variation of individual crystals from the main orientation is generally not greater than 5°. In fact the transmission pattern is very similar to that of the stretched specimens of polythene described elsewhere in this paper. Moreover, the sharpness of the rings clearly shows the crystals to be relatively much larger than those formed in a similar fashion on metals.

§ 8. POLYTHENE EMULSION

In addition to the specimens of solid polythene described above, the polymer was also available in the form of an emulsion. This consists of thick white paste, mainly polythene but containing a small percentage of other compounds necessary for stabilization.

The emulsion was diluted in about 100 times its volume of water, and several drops placed on a thin film of collodion or methyl methacrylate as a backing. When the film had dried it was examined in the electron-diffraction camera. In many films the polythene was found to coagulate in small regions, which could not be examined owing to "charging-up" of the specimen in the electron beam. Thinner films suitable for examination gave rise to patterns consisting of diffuse haloes and therefore very dissimilar from the patterns obtained from films of polythene formed by evaporation.

In a few cases some weak rings were also visible but in general these could not be reproduced consistently. These spacings were in approximate agreement with the 210, 020, 120 or 011 and 111, 201, etc., reflections. The intensities of these rings were very low, and moreover other relative intensities differed considerably from those obtained from films produced from a solvent. The intensity and infrequent occurrence of these rings shows that the quantity of crystal polythene present is very small and irregularly distributed. The haloes which form the most striking part of the diffraction pattern were found to correspond to the spacings (after application of Bragg's law) 4·55 (4·2–4·9), 2·2 (2·05–2·35), 1·2 (1·1–1·3) A. Figures in brackets indicate the approximate width of the halo.

The 4·55 halo is relatively sharp but is superimposed on a weak and broad halo with a peak intensity at about 3·6 A. At first it was not certain whether the outer haloes were in fact due to polythene or to the collodion or methyl methacrylate background. A comparison with molten polythene, described elsewhere, showed that the 2·2 and 1·2 diffractions are in fact due to polythene, but that the 3·6 halo most probably arises from the backing.

The 1·2 and 2·2 haloes coincide with the centre of a group of rings in the crystalline pattern and may be considered to derive from a blurring of the rings, due to the small size or distortion of the crystals. This explanation does not account for the 4·55 halo, which is clearly distinguishable from the 110 and 200 diffractions.

Electron-diffraction patterns of polythene emulsion show strong similarity to those of liquid polythene. The comparison will be discussed more fully in connection with the latter.

When films prepared as above are heated to 100° C. before examination, no marked changes are found in the electron-diffraction pattern, although faint rings may be observed somewhat more frequently in the unheated specimens. This change, moreover, was only observed in a few of the heated specimens, and cannot therefore be accepted as a very reliable indication of any increase in crystallinity after heat treatment.

§9. NON-ORIENTED POLYTHENE

Attempts were made to obtain films of crystalline polythene which did not show the orientation customary in films produced by evaporation on a water surface. Among the methods attempted were : heating an oriented film, etching in xylene, melting, evaporation on a thin gauze. None of these proved successful in producing a thin film of polythene with an orientation different from that described above.

§10. CONCLUSIONS

Thin films of polythene prepared by evaporation from a dilute solution of hot water are crystalline in character. The structure of these films is similar to that of the hydrocarbon described by Bunn, with the following axial spacings in the orthorhombic unit cell : $a=7{\cdot}428$ A., $b=4{\cdot}9324$ A., $c=2{\cdot}532$ A. In addition, however, the molecules are oriented by the action of the hydrophilic end-group $-\mathrm{CH}=\mathrm{CH}_2$, which can only be incorporated in the structure if the molecules are tilted, so that the *b* axis is parallel to the water surface, while the *c* axis makes an angle of 56° instead of 90° with the water surface. The crystalline character of these thin films of about 1000 A. thickness contrasts with the amorphous structure of polythene in the bulk form.

Thin films of polythene may be stretched and their altered structure examined. This is found to consist of crystals in which the *c* axis lies nearly parallel to the direction of the applied stress. The intensity distribution in the electron-diffraction pattern leads one to postulate inter-molecular forces which stabilize the crystals in their new positions, until the temperature is raised to the neighbourhood of the melting point, when the film reverts to some extent to its initial orientation.

On a metal surface, polythene molecules are found to be relatively mobile. They may be oriented mechanically to produce a structure basically similar to that of stretched films.

Polythene emulsion has a structure very different from that of the crystalline film and may, in fact, be described as of an amorphous character. There is, however, some evidence of a small proportion of crystalline polythene. The structure of the emulsion shows great similarity to that of liquid polythene.

ACKNOWLEDGMENTS

The writer wishes to thank Professor G. I. Finch, F.R.S., for advice and Dr. E. H. Rodd and I.C.I. (Dyestuffs) Ltd., for the provision of specimens and a grant during the tenure of which much of the above work was carried out.

REFERENCES

Bunn, C. W., 1939. *Trans. Faraday Soc.* **35**, 482.
Garner, W. L., Bibber, V. and King, A. H., 1931. *J. Chem. Soc.* 1533.

EFFECT OF TEMPERATURE ON THE STRUCTURE OF HIGHLY POLYMERIZED HYDROCARBONS

By A. CHARLESBY,
Imperial College of Science and Technology, London

MS. received 3 *April* 1945

ABSTRACT. An investigation into the effect of temperature on the structure of polyethylene (polythene) is made over a range extending from the temperature of liquid oxygen to temperatures at which polythene exists in the liquid condition. The following main conclusions emerge.

As with much lower polymers, the ratio a/b of the two axes increases with temperature, and tends to the value for a pseudo-hexagonal structure, but this value is never attained, since fusion sets in first.

A change from a crystalline to an amorphous pattern occurs at temperatures which are below those generally taken as the melting point, but which agree with that for which polythene transforms from a wax-like to a more transparent solid. The transition of structure extends over a number of degrees, and in this interval crystalline and amorphous patterns are superimposed.

Orientation present in crystalline films is retained when the films are heated to temperatures well above the melting point, and subsequently cooled.

A preliminary discussion is made on some of the points raised by the observations.

---

§1. INTRODUCTION

In the previous paper (Charlesby, *Proc. Phys. Soc.* **57**, 496, 1945), the structure of an ethylene polymer of high molecular weight (polythene) was studied at room temperatures by means of electron diffraction. In thin films the polymer presents a structure similar to that of the shorter hydrocarbons C_nH_{2n+2}; but, due to the enormous lengths of the molecules, the intermolecular distances in the direction of the *c* axis play an inappreciable rôle in the structure. The existence of the end group $-CH=CH_2$ in the high polymer may, under suitable conditions, modify the structure by tilting the molecules relative to the plane of the film.

The investigation described below is an extension of this work and summarizes the influence of temperature on structure down to the temperature of liquid oxygen and into the region in which polythene exists in the molten condition. Although the results obtained for liquid polythene appear to be of considerable interest, a full critical examination and discussion are postponed to a later paper.

§2. APPARATUS

Electron-diffraction patterns were taken with a Finch type of camera. Two modified types of specimen carrier were designed to permit the temperature of the specimen in the evacuated camera to be altered at will. The first type was primarily designed for temperatures below room temperature, but could be used up to about 60 or 80° c. The second enabled the specimen to be heated from room temperatures to well above the melting point of polythene.

The low-temperature design (figure 1) consists essentially of two concentric copper tubes bent upwards at one end and soldered together. The space between these communicates with the main vacuum of the camera at the other end and thus serves partially to insulate the inner tube thermally. The inner tube ends in a heavy brass disc P bearing a screwed rod.

The specimen, usually mounted on a strip of nickel gauze, is clamped to a short holder attached to this rod. It is surrounded by a shield S, which is a close-sliding fit on tube T. In this shield two apertures are provided for the incident electron beam and for the diffracted beams. Each of these apertures is surrounded by guard rings to trap water vapour which might otherwise condense on the specimen at low temperatures.

The device is carried by a brass disc sealed to the outer tubing. It is pressed inwards by the pressure of the atmosphere, and is prevented from sliding downwards by ring R. The carrier C is of the standard pattern ; by means of adjustable screws used in conjunction with springs and tombac tubing, it permits the position and orientation of the specimen to be adjusted. It is mounted against a standard port of the camera, and is held in position by a ring R′ and lug L. Rubber gaskets G, G′ are provided to increase the heat insulation of the cooling chamber T.

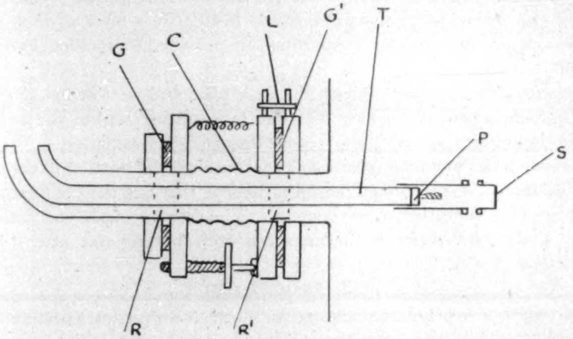


Figure 1.

In operation, the prepared specimen is fixed to the cooling chamber and the shield mounted in a suitable position over it. The complete carrier is fixed in the camera, which is then evacuated. Only when this process is completed is liquid hydrogen introduced into the cooling chamber, as this procedure reduces the possibility of ice condensing on the specimen. The electron beam is scanned over the small portion of the specimen visible through the aperture, and when a suitable adjustment is obtained the plate is exposed. Finally the vacuum is broken and a check is made to ensure that some liquid air remains in the cooling tube, thereby verifying the temperature of the specimen.

For higher temperatures the cooling chamber is filled with hot water through which steam is passed from an external boiler. The temperature is read from a thermometer placed in the water. Unfortunately this method is limited in range and the temperature readings cannot be relied upon to give the temperature of the specimen, which is fairly distant from the thermometer.

In the construction of the carrier the specimen is situated in what may be considered as a constant-temperature enclosure. Some heat losses occur, mainly by radiation, and wherever possible the apparatus is constructed of heavy metal to ensure good thermal conduction, and hence thermal equilibrium, over the surface. Some very slight temperature differences may arise on the specimen due to direct radiation to or from the camera walls through the apertures. The amount of heat supplied to the specimen by the destruction of kinetic energy of scattered electrons in the incident beam is too small to be appreciable.

Experiments with the first type of apparatus were sufficient to indicate a variation of spacings with rising temperature. The second apparatus permits a more accurate control of temperature as well as a more extensive range. It consists of a demountable furnace, electrically heated in the vacuum, and held at the end of a paxolin rod. As before, the specimen is mounted on a nickel gauze clamped to a short holder, which is screwed to a stout brass plate. The cylindrical furnace also screws on to this plate and is heated by a layer of high-resistance wire wound on asbestos paper. Apertures are provided for incident and emergent electron beams.

A thermocouple passes through the base plate and is fastened to the nickel gauze, which is bent down over it. The temperature within the enclosure is sufficiently uniform, the only appreciable loss being by radiation. Although the thermocouple was mounted several millimetres from the portion of the specimen under examination, there is no reason to believe that the temperature measured was in error by more than a few degrees. The thermocouple was standardized against a standard mercury thermometer both before and after the set of experiments.

§ 3. PROCEDURE

Thin films of A polythene, suitable for electron-diffraction examination, were prepared by evaporation of a drop of xylene solution on a hot water surface. The results are summarized in table 1. Specimens A–G were mounted on nickel gauze, I on a thin methyl methacrylate film mounted on gauze, while H and J were mounted on a very fine meshed gauze.

The change in lattice dimensions of hydrocarbons resulting from thermal expansion has been discussed fully by Müller (1930, 1941) who has shown that any increase in the c axis must be very slight, whilst an increase in both the absolute values of the a and b axes, as well as of the ratio a/b, is to be expected. In the present experiments attention has been confined to change in shape, as, in the absence of a calibration standard, absolute changes in dimensions (which are in any case small) are more difficult to determine than relative changes.

In all the measurements of crystalline polythene given in table 1 the measurements of the 200 (3·714 A.) and 020 (2·467 A.) spacings were made relative to 110 (4·109 A.), chosen as a standard. With increasing temperature the a axis is found to increase in length relative to the b axis, a result in agreement with measurements on short hydrocarbons, and which derives from a tendency to assume a pseudo-hexagonal structure at higher temperatures.

Table 1. Variation of spacing with temperature

			Calculated spacings at room temperature											
			4·109	3·714	2·467	2·253	2·055	1·926	1·73	1·61	1·27	1·21	1·098	
Specimen	Plate no.	Temperature (°c.)	Observed spacings											Accuracy
A	4921	20	4·109	3·72	2·472				1·73					±0·3 %
	4922	−183	4·109	3·67	2·485				1·74			1·225	1·10	
	4923	−183	4·109	3·73	2·503				1·75	(1·62)			1·11	
D	4929	−183	4·109	3·718	2·49			1·936	1·73	1·618		1·24	1·11	±0·5 %
	4930	65	4·109	3·745	2·44									
E	4933	72	4·109	3·761	2·472	2·249	2·07	1·925	1·724	1·599	1·257	1·21		±0·3 %
G	4974	83	4·109	3·78	2·44	2·23	2·06	1·72	1·706	1·598				±0·6 %
	4976	88	4·109	3·74	2·448	2·238								
	4977	82	4·109	3·774	2·455	2·241	2·06		1·708	1·596	1·251	1·20		±0·3 %
H	5039	20	4·109	3·71	2·46	2·25	2·07	1·94	1·72	1·605	1·255	1·20	1·10	
	5040	94	4·109	3·78	2·44									−0·5 %
	5041	111	4·52			2·12					1·24			
I	5042	107	4·42			2·17					1·21			
	5043	107	4·52			2·11			1·63 ?		1·13			
	5044	117	(4·6)			(2·1)					(1·15)			
	5045	65	4·109	3·7	2·46									
J	5059	83	4·109	3·79	2·45									
	5060	91	4·11		2·44									
			4·6			2·17						1·23		
	5059	107	4·5			2·18						1·23		
	5062	128	4·6			2·17						1·21		
	5063	78	4·5			2·13						1·23		
	5064	20	4·11	3·71	2·46 etc.									Strong orientation
	5065	105	4·7			2·21	1·23							
	5066	169	4·7			2·17	1·21							
	5067	40	4·11	3·72	2·455 etc.									

Remarks. (1) Calculated spacings at room temperature are based on the following axial spacings : $a=7{\cdot}428$, $b=4{\cdot}934$, $c=2{\cdot}532$.
(2) Where possible spacings are measured relative to 110 diffraction (4·109 A.).
(3) Halo spacings, calculated from Bragg's law, are less accurate.
(4) Specimens A–G were mounted on ordinary gauze, H and J on fine gauze, I on methyl methacrylate film.
(5) In plate 5060 strong haloes co-exist with weak rings showing orientation.

Some estimate of the accuracy of the measurements given in table 1 was obtained by measuring several of the clearer patterns on two or three occasions; the extent of agreement served as a measure of the accuracy involved.

The change in shape is best traced from the change in the ratio *a*/*b*, and is relatively greater over the temperature range 20–80° than it is from −183° c. to 200° c. This is to be anticipated from the rapid rise in the repulsive forces between two approaching atoms. As the amplitude of thermal vibrations increases, the skew form of the negative potential-energy curve results in an increased average distance between atoms as shown in figure 2.

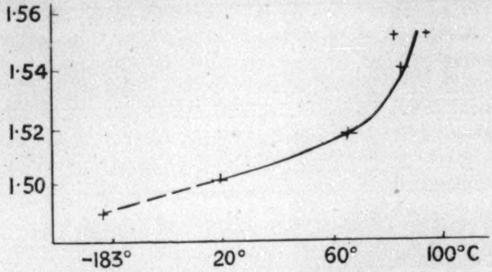


Figure 2. Variation of *a*/*b* with temperature.

§ 4. CHANGES IN STRUCTURE

At temperatures of about 90° c. a profound change in structure occurs, which is evidenced by the transformation from a ring to a halo pattern. The centres of the haloes correspond to spacings of about 4·6, 2·15 and 1·22 A., as calculated from Bragg's law. These must be identical with the diffuse band at 4·5 to 4·6 observed for liquid hydrocarbons (e.g. Müller, 1932), so that at temperatures well below the point generally accepted as the melting point the structure of polythene must be that of a liquid. It is known that at temperatures of this order polythene assumes a more glass-like appearance, and the change in transparency must be associated with the more amorphous character of the pattern obtained. It is of interest to note that many other transparent polymers, e.g. methyl methacrylate, give rise to electron-diffraction patterns which are very similar in their general appearance, but differ somewhat in the spacings, relative intensities and width of the innermost haloes.

Müller (1932) has shown that as the temperature is raised the structures of hydrocarbons $C_{21}H_{44}$ to $C_{27}H_{56}$ assume a state of pseudo-hexagonal symmetry before reaching the liquid state, but this is no longer true for $C_{29}H_{60}$ and longer molecules, in which the liquid state is first attained. In the hexagonal structure three adjacent molecules define an equilateral triangle ($\phi = 60°$); in polythene the liquid state is reached when ϕ is about 57°.

Before the hexagonal structure is attained, hydrocarbons $C_{24}H_{50}$ and above pass through a transition point, in which the structure shows an abrupt change in dimensions. In $C_{29}H_{60}$, for example, the length of the *a* axis increases by

5% within a fraction of a degree. The expansion of the polythene lattice has not been traced very closely, but an expansion of this amount does not occur below the melting point. The absence of a transition point for polythene is in any case to be anticipated, since already for $C_{44}H_{90}$ transition and melting points almost coincide.

Of considerable interest is the occurrence together of both halo and ring pattern (e.g. specimen J, plate 5060). This coincidence of the two types of structure may be detected on a number of plates in the neighbourhood of 85°. Over a temperature range of perhaps 10° amorphous and crystalline structure may co-exist in a small portion ($<10^{-2}$ mm^2) of a thin film.

This extended range of melting is a strong argument in favour of that theory of the structure of high polymers which envisages the solid as consisting of a number of crystalline regions, held together by molecules common to the structure of several of these regions. Since the forces of cohesion of these molecules must generally be lower than those situated in the bulk of the crystals, the former will first be freed as the temperature approaches the melting point, and form an amorphous or liquid system. The more perfect crystalline regions will continue to exist in equilibrium with the liquid (i.e. amorphous structure) until higher temperatures are reached.

§5. COMPARISON WITH HYDROCARBONS OF LIMITED LENGTH

In the absence of absolute measurements of the axial dimensions of polythene at different temperatures, no direct comparison can be made between its expansion and that of shorter hydrocarbons such as $C_{23}H_{48}$ and $C_{29}H_{60}$, studied by Müller. A comparison of the shape of the cross-sectional area may, however, be made in terms of the axial ratio a/b at various temperatures. This is shown in table 2 at room temperatures, at temperatures a few degrees below the melting point, and, in the case of $C_{29}H_{60}$, immediately below and above the transition point.

Table 2. Variation of a/b at various temperatures

	Liquid air or oxygen	Room temperature	Transition point		Below melting point
Polythene	1·49	1·505			1·55
$C_{29}H_{60}$	1·47	1·502	1·513	1·601	1·662
$C_{23}H_{48}$		1·496			1·73

Note : 1·73 corresponds to a hexagonal structure.

It is obvious that the change in dimensions becomes progressively smaller in the higher hydrocarbons, in spite of the more extensive temperature range available.

The comparison is of interest from the point of view of the character of the thermal vibrations of the hydrocarbon chain. In the past it has been suggested that these vibrations occur mainly by a bodily rotation of molecules about their axis. Further work (e.g. Müller, 1940) has shown that tensional vibrations of the molecular chains must play an important part. In the polythene the length of the molecules precludes the possibility of appreciable bodily rotation. The magnitude of the torsional vibration may to some extent be gauged from the deformation of the structure from its initial dimensions at liquid-air temperatures.

§6. PERSISTENCE IN ORIENTATION

All the films studied were produced by evaporation on the water surface and showed the {20} form of orientation produced under these circumstances. This orientation is readily detectable by a study of the relative intensities of the diffraction rings, especially if the film is not normal to the incident electron beam. Present experiments permit an investigation into the extent to which this orientation is conserved at higher temperatures.

In tilted specimens the initial orientation present may be best judged from the arcing of the 020 reflection, which is most intense in a direction parallel to the axis of rotation of the specimen. In all halo patterns no orientation is detectable, even in tilted specimens.

At low temperatures (– 183° C.) the orientation pre-existing at room temperature remains unaltered. As the temperature is raised to 80° C., there is no appreciable change in the orientation, the 020 remaining arced, although there is an overall reduction in the intensity of the reflections relative to the 110.

In specimen I an examination of the structure was made after the specimen had remained in the amorphous condition (corresponding to a liquid structure) for a period of nearly one hour. Some orientation was found to persist on cooling to 65°, even though the specimen had previously reached a temperature of 117°. In this specimen the polythene film was spread on the surface of methyl methacrylate, and the presence of the latter might conceivably cause a re-orientation on cooling. Specimen J was therefore prepared on a support consisting of a fine nickel gauze, and was heated to 128° C. The orientation of this film after cooling was most marked. Finally a temperature of 168° was reached, but even after this treatment the remaining orientation was still strong.

It may be suggested that re-orientation occurs subsequently to cooling under the influence of the gauze support or of surface tension. Neither of these suggestions can, however, account for the type of tilted orientation actually observed, which is typical of films oriented under the action of the water surface. In a globule of molten polythene in which the molecules are arranged at random no forces can exist which, when the film is cooled, re-orient the molecules at an angle to the surface similar to that produced on a water surface.

The important conclusion is therefore reached that in thin polythene films, and probably in all molten high polymers, the initial structure of the solid is to a considerable extent preserved in the molten form for long periods. The oriented molecules may then serve as a nucleus of condensation in the subsequent process of solidification.

The present experiments give strong evidence in favour of the hypothesis that crystal structure is to some extent retained in the liquid form.

§7. STRUCTURE OF LIQUID POLYTHENE

X-ray diffraction patterns of hydrocarbons consist of a diffuse halo, the spacing of which corresponds to 4·5 to 4·6 A. if Bragg's law is used to translate distances measured on the x-ray plate into spacing between planes. The haloes obtained from polythene films at temperatures of 90° C., and which persist at least up to 168° C., must be of a similar origin. In addition to the 4·6 A. halo, haloes corresponding to spacings 2·15 and 1·22 A. are observed in the electron-diffraction patterns. This increase in the data provided renders a search for a suitable structure more hopeful.

The 2·15 A. halo occurs at a spacing overlapping reflections such as 111, 201, etc., and its presence may be explained as deriving from these reflections in a distorted or minute structure. Moreover, the 1·22 halo almost certainly corresponds to the 002 reflection, with, perhaps, an overlap from such reflections as 202, etc. The origin of the 4·6 halo nevertheless remains obscure.

The internal structure of liquid hydrocarbons is far from clear. If it partakes of a crystalline character, the minute size of the crystals accounts for the diffuse character of the reflections observed. The application of Bragg's law is then at least approximately justifiable, but the position of the haloes cannot be readily explained. On the other hand, if the structure is essentially amorphous in character, the application of Bragg's law requires further justification.

§8. POLYTHENE IN THE FORM OF A LIQUID AND EMULSION

The pattern of polythene emulsion at room temperature shows a remarkable similarity to those of molten films of polythene, as is shown in table 3.

This great similarity indicates that the intermolecular spacings and orientation must be similar, and hence that polythene emulsion exists as a supercooled form of liquid polythene.

A further specimen of polythene containing a very small percentage of extraneous matter was examined after standing for about one year. The patterns obtained were more nearly crystalline with spacings corresponding to those of crystalline films. The sharpness of the rings showed that the crystals could not be very extensive. Moreover, the intensity distribution was such that the orientation of the films produced by evaporation on a water surface no longer existed. In most cases only the 110 reflection was observed.

Table 3

		A.	A.		
Molten polythene	Average of 10 plates	4·6	2·15	1·22	Centre of haloes
		4·1–5·0	2·05–2·25	1·10 1·35	Width of haloes
Emulsion	Average of 2 plates	4·6 (3·6)	2·2	1·2	Centre of haloes

§9. CONCLUSION

As the temperature of a film of polythene is raised, its structure is modified by the tendency to form a pseudo-hexagonal structure. This is not, however, achieved, as melting sets in at a lower temperature.

Polythene in the molten condition presents an amorphous structure, represented in the electron-diffraction pattern by three haloes at 4·6, 2·15 and 1·22 A., that at 4·6 A. being the most sharp and intense. Over a range of several degrees the amorphous and crystalline types of structure co-exist throughout the film. That the melting point of polythene is not sharp is common knowledge, but it is of interest to find this associated with a gradual transition from the fully crystalline to a fully amorphous structure. In shorter hydrocarbons the transition is far more sudden. It may be argued that molecules of varying lengths form corresponding crystals with varying melting points, but this cannot be supported by the evidence, since the temperature region in which the transition from crystalline to amorphous structure occurs is far lower than that to be anticipated for crystals of lengths approximating to those of the highly polymerized molecules. If the change in pattern is due to breakdown of crystal structure, the individual crystals must be far shorter than the individual molecule. Thus, at room temperature, molecules must enter into the structure of a number of these crystals.

It is then reasonable to assume that some molecules are more loosely held, and on being released from many of their bonds at temperatures below the melting point of the bulk crystal form the amorphous phase. As the temperature is further raised, the proportion of crystalline structure rapidly decreases.

The region in which the transition from crystalline to amorphous structure occurs extends over a number of degrees, but, certainly as far as can be judged from the electron-diffraction pattern, the change is complete at temperatures below the melting point as the latter is generally understood, and, moreover, appears to coincide with the usual transition to a more glass-like material.

The conservation of the orientation initially present in the films even when these are heated to 168° C. shows that the molecular arrangement is not entirely random even at temperatures well above the melting point. Further work is clearly desirable on the rate at which the initial orientation is lost when the film is held in the molten state for varying periods. These considerations lead one to envisage a statistical fluctuation of temporary crystallinity over small regions, sufficient to ensure mobility above the melting point, over macroscopic periods of time, yet serving to retain some "memory" of initial structure.

REFERENCES

CHARLESBY, A., 1945. *Proc. Phys. Soc.* **57**, 496.
MÜLLER, A., 1930. *Proc. Roy. Soc.* A, **127**, 417 ; 1932. *Ibid.* **138**, 514 ; 1940. *Ibid.* **174**, 137. 1941, *Ibid.* **178**, 232.

NOTE ON DIFFUSION IN THE IONOSPHERE

By J. C. JAEGER,
The University of Tasmania

MS. received 31 *May* 1945

ABSTRACT. The equation of diffusion for a one-dimensional region in which the coefficient of diffusion varies exponentially is solved, and the results applied to the diffusion of ions in the ionosphere. Numerical results are given for three special initial distributions of ion density.

THE effect of diffusion of ions in the ionosphere has been very little studied. The general equations, including diffusion, recombination and ion production, were set up by Hulburt (1928) but are quite insoluble. Recently Bagge (1943) has shown, by numerical integration in a special case, that diffusion will greatly reduce the ion density in the upper parts of a Chapman region, although it leaves that near the maximum unaltered.

The object of this note is to set out briefly the results for one case in which the exact solution is possible, namely, that of an atmosphere in which the particle density varies exponentially with height, and in which the effects of recombination and the earth's magnetic field are neglected. This allows an estimate of the effect of diffusion to be made which may be useful in the higher regions, where recombination is of less importance. The neglect of the earth's magnetic field is, of course, serious ; its effect is considered briefly later.

Let N be the number of ions per c.c. at height z, and let $1/b$, supposed constant, be the scale height of the atmosphere, so that the coefficient of diffusion at height z, which is inversely proportional to the particle density at this height, may be written

$$\kappa e^{bz}, \qquad \ldots\ldots(1)$$

where κ is a constant. Then if $I(z, t)$ is the rate of ion production at height z at time t, the differential equation satisfied by N is [cf. Hulburt (1929) or Bagge (1943)]

$$\kappa e^{bz}\left(\frac{\partial^2 N}{\partial z^2} + 2b\frac{\partial N}{\partial z} + b^2 N\right) - \frac{\partial N}{\partial t} + I(z, t) = 0. \qquad \ldots\ldots(2)$$

Since in practice we are only interested in large values of z, and those ions which diffuse to smaller values of z will be removed by recombination, it is sufficient to consider (2) in the region $-\infty < z < \infty$ with the boundary conditions N finite as $z \to \pm\infty$. We proceed to solve (2) in this region for $t > 0$, with a given initial value $N_0(z)$ of N when $t = 0$, for the case* $I(z, t) = 0$.

---

* There is no additional difficulty in including ion production, but it is perhaps pushing a simple model too far.

Writing

$$y=Ne^{bz}, \qquad x=e^{-bz}, \qquad \ldots\ldots (3)$$

(2) becomes in the case $I(z,t)=0$,

$$\frac{\partial}{\partial x}\left(x\frac{\partial y}{\partial x}\right)-\frac{1}{\kappa b^2}\frac{\partial y}{\partial t}=0, \qquad 0<x<\infty, \qquad t>0, \qquad \ldots\ldots (4)$$

to be solved with

$$y \text{ finite as } x\to 0, \text{ and as } x\to\infty, \qquad \ldots\ldots (5)$$

and with

$$y\to y_0(x)=\frac{1}{x}N_0\left(\frac{1}{b}\ln\frac{1}{x}\right), \quad \text{as} \quad t\to 0. \qquad \ldots\ldots (6)$$

The solution* of (4), (5) and (6) is

$$y=\frac{1}{\kappa b^2 t}\int_0^\infty y_0(u)\, I_0\left(\frac{2\sqrt{(ux)}}{\kappa b^2 t}\right) e^{-(x+u)/\kappa b^2 t}\, du. \qquad \ldots\ldots (7)$$

We shall write k for the dimensionless quantity:

$$k=\kappa b^2 t. \qquad \ldots\ldots (8)$$

It is easy to verify that (7) satisfies the differential equation and initial and boundary conditions.

To illustrate the behaviour of the solutions of (4) we consider two cases in which (6) takes a specially simple form.

Case (i). $N_0(z)=n_0\delta(z-z_0)$, where δ is the Dirac δ function, that is, a concentration of n_0 ions is released at $t=0$ at the level z_0. In this case the solution is

$$\frac{N}{bn_0}=\frac{x}{k}I_0\left(\frac{2\sqrt{(xx_0)}}{k}\right)e^{-(x+x_0)/k}, \qquad \ldots\ldots (9)$$

where $x_0=e^{-bz_0}$.

The right-hand side of (9) as a function of x involves only the quantity k, but if we express it in terms of z, b appears as well. As the results are a little more physical when expressed in terms of z, they are plotted in this way in figures 1 and 2 for the value $b=2\times 10^{-7}$ corresponding to a scale height of 50 km., that usually observed in the F region. Figure 1 shows the diffusion of a concentration of n_0 ions initially at $z_0=200$ km., and figure 2 that of a concentration initially at $z_0=400$ km., for various values of $k=\kappa b^2 t$. A large downward motion is at once apparent. Whether this is of physical importance depends on whether values of k of the order of those for which the curves are drawn can be attained in practice. The coefficient of diffusion at any height is $0{\cdot}41\gamma u$, where γ is the mean free path at that height and u the mean molecular velocity. Taking the value of γ at 250 km. to be 10^6 cm. (Hulburt (1937), table 1) and for u its value at 220° K., 10^5 cm. per sec., we get $4{\cdot}1\times 10^{10}$ for the value of the coefficient of diffusion.

---

* For the method of solution using the Laplace transformation, see Carslaw and Jaeger, *Operational Methods in Applied Mathematics*, § 68 (Oxford, 1941). $I_0(x)$ is the modified Bessel function of the first kind of order zero.

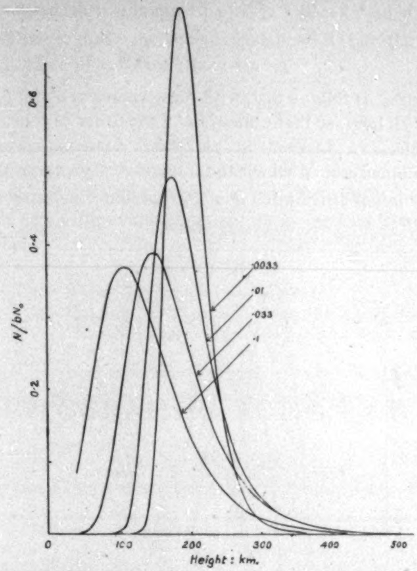


Figure 1. **Diffusion of a concentration of ions initially at 200 km. Scale height 50 km. The numbers on the curves are the values of *k*.**

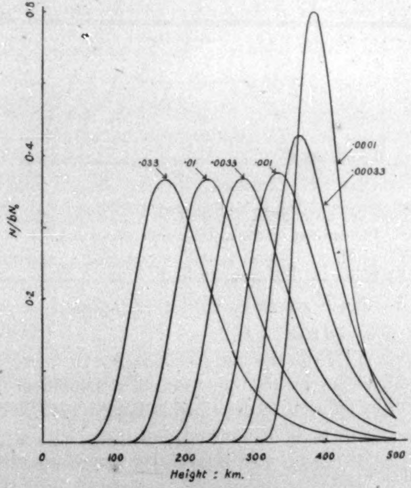


Figure 2. **Diffusion of a concentration of ions initially at 400 km. Scale height 50 km. The numbers on the curves are the values of *k*.**

If we give b the value* 2×10^{-7} obtained from the scale height, 50 km., of the F region, we have from (1), for $z=2{\cdot}5\times10^7$,

$$\kappa e^5=4{\cdot}1\times10^{10},$$

that is, $\kappa=2{\cdot}8\times10^8$. It follows that, with these values of κ and b, k has the value 0·04 at the end of an hour, so that values of k of the order of those in figures 1 to 3 seem quite possible.

Case (ii). Another case in which the integral (7) can be evaluated simply is that in which the initial distribution is of the familiar Chapman type, say

$$N_0(x)=N'x^{\frac{1}{2}}e^{-cx}, \qquad \ldots\ldots(10)$$

where x is given by (3). In this case we find

$$N=N'x(\pi/kk')^{\frac{1}{2}}e^{-x(2k'-1)/2kk'}I_0(x/2kk'), \qquad \ldots\ldots(11)$$

where k is defined in (8) and $k'=1+kc$.

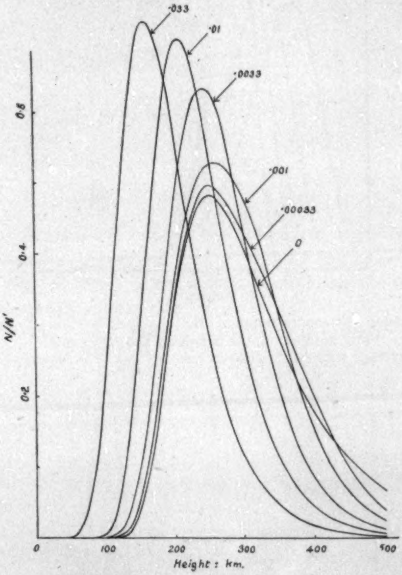


Figure 3. Diffusion of a distribution of ions which is initially of Chapman type. The numbers on the curves are the values of k.

In figure 3 curves of (11) for the case $c=80$, $b=2\times10^{-7}$, and various values of k are given. The tendency for the upper ions of a theoretical Chapman region to move downwards by diffusion as described in Bagge (1943) is clearly shown.

---

* If values of b based on the kinetic theory are used, for example those deduced from Hulburt's (1937) table 1, much smaller values of k are obtained, but these do not allow for the probable high temperature at these levels.

The effect of the earth's magnetic field is to leave the coefficient of diffusion unaltered for diffusion along the magnetic field, but, for diffusion across the magnetic field, in place of κe^{bz}, we have

$$\frac{\kappa e^{bz}}{1+ae^{2bz}},$$

where a is a known constant depending on the field. The resulting equation is not soluble in terms of known functions. In temperate latitudes the effects described above will not be altered greatly, but at the magnetic equator vertical diffusion will not be an important agent for removing ions from high levels.

REFERENCES

Bagge, E., 1943. *Phys. Z.* **44**, 163.
Hulburt, E. O., 1928. *Phys. Rev.* (2) **31**, 1018.
Hulburt, E. O., 1937. *Rev. Mod. Phys.* **9**, 44.

THE PHOTOGRAPHIC ACTION OF X RAYS

By S. R. PELC,
Physics Department, Hammersmith Hospital, London

Communicated by Dr. L. H. Clark, 20 *May* 1945

ABSTRACT. An account of theoretical considerations of the photographic action of x rays, based on the absorption process of x rays in photographic emulsions, and found to be in good agreement with experimental results, is presented in this paper. An equation giving the variation of the photographic action with wave-length is deduced. Consideration is given to the number of grains made developable per absorbed photon. The conclusion is reached that the energy required to make one grain developable is, for a given emulsion, constant for effective wave-length from 47 x.u. to 395 x.u.

§1. INTRODUCTION

The results of previous investigations show that the shape of the density-exposure curve is essentially the same for different qualities of radiation (Glocker and Traub, 1921). Berthold (1925) found that in order to produce the same density with x-ray beams of different penetrating power, more energy must be absorbed in the emulsion from a penetrating radiation than from a less penetrating one. Similar investigations (G. E. Bell, 1936 ; E. E. Smith, 1943) with heterogeneous x rays show that the exposure, expressed in roentgens, required to produce a certain density passes through a minimum in the region of 0·25 mm. copper H.V.L.

M. Blau and K. Altenburger (1922) and L. Silberstein (1922) assumed that at any stage of the exposure the number of affected grains (n) is proportional to the number of unaffected gains, i. e.

$$dn=(N-n)\epsilon\, dI, \qquad \ldots\ldots(1\,a)$$

where N is the total number of grains per unit area, ϵ is an arbitrary factor and I the exposure.

On integration, we find

$$n = N(1 - e^{-\epsilon I}). \qquad \ldots\ldots(1\,b)$$

L. Silberstein (1922) regarded each grain as presenting to the photons a target of area *a*, and gave equation (1 *b*) in the form

$$n = N(1 - e^{-a\beta Q}) \qquad \ldots\ldots(1\,c)$$

where Q is the number of incident photons and β an arbitrary factor. The meaning of β has been differently interpreted : by L. Silberstein and A. P. H. Trivelli as the fraction of the area of a grain which acts as an efficient target for light quanta, and by A. Charlesby as the probability of a grain becoming developable when struck by a photon. The density is taken as proportional to the number of grains, and it was found that the density-exposure curves for x-ray exposures can be well represented for moderate densities by an equation of the form of equation (1 *b*) (Bell, 1936 ; Charlesby, 1940). It can be shown that the speed of a photographic emulsion determined in the usual way from the curve of density against log (exposure) is proportional to ϵ whenever the density-exposure curve can be represented by (1 *b*) ; ϵ is sometimes referred to as the *speed* in this paper. The following considerations show that the arbitrary factor ϵ can be evaluated from the absorption process of x rays in the silver halide of the emulsion and the known number of incident photons per unit area, when the exposure is expressed in international roentgens.

§ 2. THEORETICAL

The number of grains affected, n, will be proportional to the number of quanta absorbed, and to the number of grains which each absorbed quantum renders developable, as well as to the probability of a hitherto unaffected grain being affected. Taking q to be the number of photons absorbed and η the number of grains affected per quantum absorbed, then

$$dn = \frac{N-n}{N}\,\eta\,dq. \qquad \ldots\ldots(2)$$

From the known absorption of x rays ($I = I_0 e^{-\mu d}$), the number of quanta absorbed is found to be

$$dq = (1 - e^{-\frac{\mu}{\rho}g})\,dQ, \qquad \ldots\ldots(3)$$

where Q is the number of incident quanta, μ/ρ the mass-absorption coefficient of silver bromide or silver chloride, and g the total amount of silver halide in grams per sq. cm. of emulsion. The amount of silver in the emulsion is of the order of 10^{-4} per sq. cm., so that for the qualities of x rays used in this investigation, (3) can be simplified to (3 *a*) with an error not exceeding 1% :

$$dq = \frac{\mu}{\rho}\,g\,dQ. \qquad \ldots\ldots(3\,a)$$

The total amount of silver bromide is $N\gamma$, where γ is the mass of the undeveloped grain, and substituting (3*a*) in (2) gives

$$dn = (N-n)\eta\,\frac{\mu}{\rho}\,\gamma\,dQ. \qquad \ldots\ldots(4)$$

For x rays measured in international roentgens (r.) the number of quanta per sq. cm. cross-section of the beam can be calculated from the energy necessary to liberate one pair of ions in air (35·5 ev.) and from the definition of the roentgen, and is found to be

$$Q=\frac{\lambda r}{(\tau+\sigma_a)_{\mathrm{air}}}\times 0{\cdot}597\times 10^7, \qquad \ldots\ldots(5)$$

where λ is the wave-length in Angstrom units, τ the photoelectric absorption, and σ_a Klein-Nishina's scatter-absorption coefficient in air.

Substituting the value for the number of incident quanta (5) in equation (4) gives the final expression,

$$dn=(N-n)A\eta\gamma\,dr, \qquad \ldots\ldots(6)$$

where

$$A=\frac{\mu_{\mathrm{AgBr}}\lambda}{(\tau+\sigma_a)_{\mathrm{air}}}\times 0{\cdot}923\times 10^6. \qquad \ldots\ldots(6\,a)$$

Integration from 0 to r yields

$$n=N(1-e^{-A\eta\gamma r}). \qquad \ldots\ldots(7)$$

Comparing equation (7) with equation (1 *b*) we see that the exponential factor ϵ is now expressed as a function of separately assessable terms and elementary constants when the exposure is expressed in international roentgens:

$$\epsilon=A\eta\gamma. \qquad \ldots\ldots(7\,a)$$

The proportionality between the speed (ϵ) and the average mass per grain as indicated in equation (7 *a*) is obviously in contradiction to the work of L. Silberstein and A. P. H. Trivelli (1930) and of A. Charlesby (1940), and will be referred to in more detail in the discussion of results.

For photographic work, the measurement of densities is, of course, preferable to counting the number of grains, and a few remarks about the connection between the number of affected grains, development and the resulting density may be inserted here. In general only a fraction α of the affected grains will be developed, and equation (7) can be written

$$n_d=\alpha n=\alpha N(1-e^{-\epsilon r}). \qquad \ldots\ldots(7\,b)$$

The fraction α can be calculated from the increase of density with time of development.

From a consideration of the combined transparency of a number of layers in a photographic emulsion, P. G. Nutting (1913) deduced the formula

$$d=-p\log\left(1-\frac{S}{p}\right), \qquad \ldots\ldots(8)$$

where d is the observed density, p the number of layers, defined as the thickness of the emulsion divided by the average thickness per grain, and S the total projected area of the developed grains. If S/p is small compared with p, equation (8) can be approximated by Nutting's simplified formula,

$$d=0{\cdot}4343S. \qquad \ldots\ldots(9)$$

The number of grains affected, as expressed in equation (7), and the corresponding density as expressed by (8) in connection with (7), will be strictly valid for emulsions consisting of grains of equal size and approximately valid for emulsions with relatively small variation in grain-size.

For measurements in the region of low densities, when the density can be well approximated by equation (9), equations (7 *b*) and (9) combined give

$$d=0{\cdot}4343\alpha\bar{a}N(1-e^{-\epsilon r}), \qquad \ldots\ldots(10)$$

where $\bar{a}$ is the average projected area per grain. For small values of the term in the exponent, and low densities, i.e. as far as the familiar linear portion of the density-exposure curve is in the region of low densities, equation (10) can be further simplified to

$$d=0{\cdot}4343\alpha\bar{a}NA\eta\gamma r. \qquad \ldots\ldots(11)$$

Substituting $g=N\gamma$ as before gives the most convenient expression for low densities in slow emulsions of high contrast, namely,

$$\frac{d}{r}=0{\cdot}4343\alpha\bar{a}A\eta g. \qquad \ldots\ldots(11\,a)$$

The error incurred in these approximations calculated for the emulsions used in this investigation from the known difference between equations (8) and (9) and between equations (10) and (11) respectively is up to 4% for one of the emulsions chosen for experimental investigation (film Y) and below 1% for film X for densities smaller than 0·5 ; the simplified equations were therefore used for film X only.

§ 3. EXPERIMENTAL

The equations resulting from the theoretical considerations of the first part contain, with the exception of the quantum efficiency η, factors which can be either calculated or determined without recourse to X-ray exposures. To test the validity of the theory an experimental investigation into the dependence of the photographic action of X rays on wave-length, and a comparison of measured densities with calculated values was undertaken.

Two types of commercial film with widely different characteristics, as shown in table 1, were chosen for experimental investigation.* The films were exposed to X rays in a special cassette made from $^1/_{16}$-inch brass sheet, which permitted the exposure of a strip of 2 cm. and accommodated up to 25 mm. thickness of lead above the part of the film not intended for exposure. The film was enclosed in thin black paper, backed by about $^1/_{16}$ inch of cardboard and slipped into a black envelope, similar to the usual double wrapping. No influence from backscatter from the backing materials could be observed. The exposures were controlled by ionometric measurements, for which instruments calibrated at

Table 1

Film	Total amount of silver bromide (gm. per sq. cm.)	Average diameter of the developed grains (cm.)
X	$6{\cdot}2\times10^{-4}$	$0{\cdot}8\times10^{-4}$
Y	$2{\cdot}6\times10^{-4}$	$3{\cdot}0\times10^{-4}$

---

* Thanks are due to Mr. R. Tupper, of the Department of Chemistry, British Postgraduate Medical School, who determined the amounts of silver per unit area quoted in tables 1 and 3.

the N.P.L. against the standard free-air chamber were used. The error of the ionometric measurements was ± 3% for radiations up to H.V.L. = 2 mm. Cu, and is unlikely to exceed ± 5% for the harder radiations. The radiation from three X-ray tubes, excited by constant-potential generators, covering a range from 90 to 500 kv., were used in order to obtain X-ray beams within a wide range of penetrating powers for the exposures ; the target-film distances were varied between 60 and 150 cm. to allow suitable times of exposure. In some of the series, standard tints exposed on each film served to eliminate the possible influence of differences in development, and in all cases all films belonging to one series of experiments were developed in the same developer within as short a time as possible. Specular densities measured with a microphotometer and barrier-layer cell calibrated against a step-wedge could be reproduced within ± 2%.

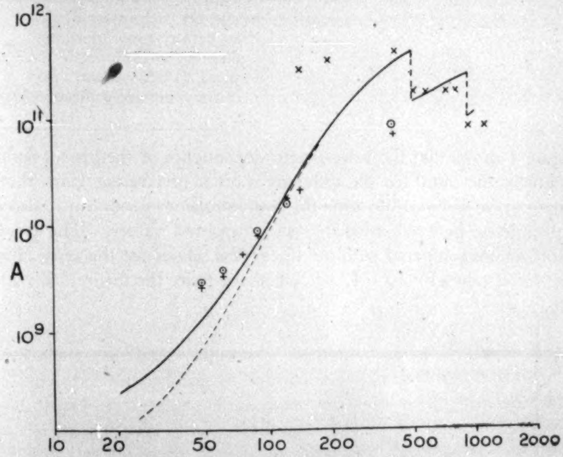


Figure 1.

As the silver in the emulsion is the dominant factor in the absorption process which leads to the photographic action, the determination of the effective wave-length of the heterogeneous X-ray beams used was based on absorption measurements in silver in preference to the more usual copper. Comparative measurements showed variations in the measured effective wave-length of about 3 X.U. for the hardest radiations, rising to 10 X.U. for the softest radiation used. The absorption coefficients used were derived from various published data (Compton and Allison, 1935 ; Jones, 1936 ; Victoreen, 1943).

Figure 1 shows calculated values for A as defined by equation (6 *a*) for wave-length of 20 to 1000 X.U. plotted on logarithmic scales ; the broken line refers to values of A when $(\tau + \sigma_a)_{AgBr}$ instead of μ_{AgBr} is used in equation (6 *a*). According to equation (7 *a*) the speed, and similarly, according to equation (11 *a*), the density per roentgen on the linear portion of the density-exposure curve would be proportional to A for radiation of different penetrating powers if η were independent of the effective wave-length, the other factors being naturally unaffected by the wave-length. The experimental results of exposures to x rays from $\lambda = 47$ X.U. to 395 X.U. are included in figure 1, and represent values for the speed for film Y and values for the density per roentgen for film X (see table 2). Berthold's values (1925) for the density divided by arbitrary units of ionization measured with monochromatic beams are included to extend the comparison into the region of the K-absorption edges of silver and bromine. All values were arbitrarily matched to the curve at one point.

Table 2

Film	Marked in figures 1 and 3	Experimental values in terms of
X	+	density per r. (constant *d*)
Y	⊙	ϵ from density-exposure curves corrected by (8)
—	×	Berthold (1925) density per arbitrary unit of ionization

Figure 1 shows that the wave-length dependence of the photographic action is essentially the same for the different ways of derivation from the density-exposure curves and also the same for both emulsions used, but it also reveals a wide divergence between calculated and measured values. The trend of the measured values compared with the theoretical values for the term *A* is unmistakable. The values for $\lambda_{\text{eff.}} = 47$ X.U. are about twice the theoretical value, whilst the values for $\lambda_{\text{eff.}} = 395$ X.U. are only a quarter, with a definite trend for the intermediate values, differences which are clearly far beyond the possible error estimated above. Berthold's values show the same tendency in the region of short wave-lengths as well as between the absorption edges.

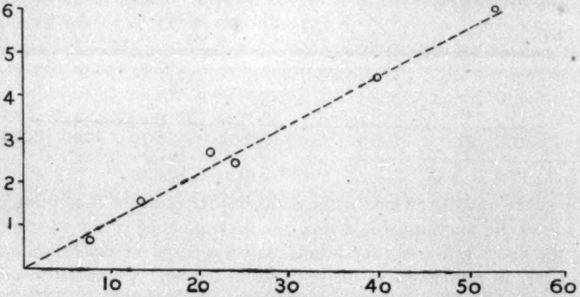


Figure 2.

Besides the factor *A*, the only parameter in equations (7 *a*) and (11 *a*) which may depend on the quality of the radiation used is the quantum efficiency, which was therefore investigated by calculating the number of photons absorbed and counting the number of developed grains per unit area under the microscope. Equation (11 *a*) was used to calculate the quantum efficiency from observed densities. Both methods give essentially the same results.

Figure 2 shows for film X the number of grains actually counted per absorbed photon against the "effective" frequency as a measure of the average energy of the absorbed x-ray quanta. Figure 2 suggests, as indicated by the dotted line, a linear increase of the quantum-efficiency with the frequency. This means, however, as comparison with equation (2) shows, that the number of grains made developable is proportional to the absorbed energy, independent of the wavelength. The elementary processes leading to these observed quantum efficiencies are not easily assessed, and no theory is attempted at this stage.

Using the Duane-Hunt relation, the observations can conveniently be presented in the form of equation (12), where k is the number of grains affected per kev. of absorbed energy, and, therefore, the reciprocal of k is the energy in kev. necessary to make one grain developable :

$$\eta = 12{\cdot}35k/\lambda. \qquad \ldots\ldots(12)$$

Inserting (12) in (7) gives

$$\epsilon = Bk\gamma, \qquad \ldots\ldots(13)$$

where

$$B = \frac{\mu_{\mathrm{AgBr}}}{(\tau + \sigma_a)_{\mathrm{air}}} 1{\cdot}14 \times 10^7 \qquad \ldots\ldots(14)$$

and

$$n = N(1 - e^{-Bk\gamma r}). \qquad \ldots\ldots(15)$$

For the linear part of the density-exposure curve we may write

$$\frac{d}{r} = 0{\cdot}4343\alpha\bar{a}gBk. \qquad \ldots\ldots(16)$$

Figure 3 gives calculated values for B, plotted as $\log B$ against $\log \lambda$. It should be noted that the ordinate in figure 3 is drawn to twice the scale used in figure 1, and, therefore, the slope of the experimental values appears to be steeper. The fall in the efficiency for harder radiations found by Berthold does not appear in figure 3, and was obviously due to the values for the absorption coefficients used by Berthold in 1925. The agreement between experimental and calculated values is satisfactory for all series, which are matched to the curve as in figure 1 by using one free parameter, the calculation of which seems desirable and can be attempted by the use of equations (13), (15) and (16). A discussion of the different factors involved may help to assess the accuracy to be expected for the comparison between measured and calculated values in tables 3, 4 and 5.

Determination of the factor k involves the measurement of the quantum efficiency at a certain stage of development, and has to be extrapolated for maximum development, because equation (13) was deduced from a consideration of the absorption in all grains and must, therefore, be referred back to full utilization of all affected grains. Term B includes the absorption coefficients of silver bromide and air. The influence of the divergencies in the various published absorption coefficients has to a certain extent been eliminated by the choice of silver for the absorption measurements in the present investigation ; the absorption coefficients for air are calculated values for the whole range. The photometric constant defined as the amount of silver per unit area divided by the density can be calculated from the value for the average mass per grain (γ), involving, apart from the measurement of density and mass of silver per unit area, the estimation of the average area per grain based on the microscopically measured average diameter per grain, i. e. the square of the diameter and the shape of the grain will influence the result. The composition of the emulsions used from various classes of grain might to a certain extent influence the calculation from basic data, although it does not influence the arbitrary values in figure 3.

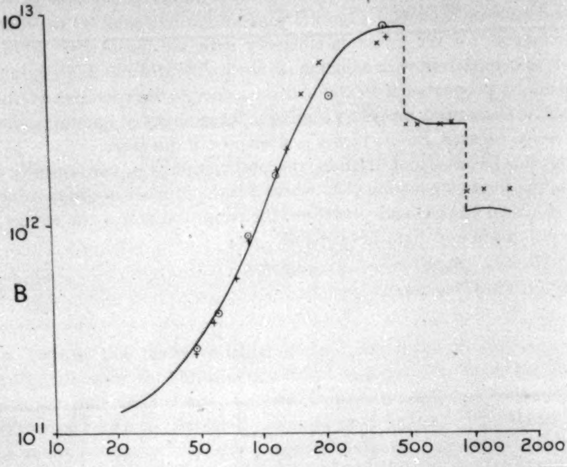


Figure 3.

Table 3

Film	λ X.U.	ϵ meas.	η meas.	k grains/ kev.	γ calc. gm./grain	Photometric constant calculated (gm./dm.^{2})	Photometric constant measured (gm./dm.^{2})
X	395	0·03	2·3	0·073	$4{\cdot}8\times10^{-14}$	$1{\cdot}0\times10^{-3}$	$1{\cdot}1\times10^{-3}$
Y	87	0·73	68	0·48	$1{\cdot}9\times10^{-12}$	$2{\cdot}8\times10^{-3}$	$3{\cdot}3\times10^{-3}$

Table 3 gives, for the two emulsions used, measured values for the factor ϵ and for the quantum efficiency (η), the latter being corrected for maximum development. The values for the number of grains made developable per kev. absorbed are calculated from equation (12) using the corresponding values for the quantum efficiency. The average mass of AgBr per grain (γ) calculated from equation (13) was used to calculate the photometric constant in grams of silver per sq. dm., no form factor (Arens, Eggert and Heisenberg, 1931) for the influence of the shape of the grains being taken into account. The agreement between calculated and measured values for the photometric constant can be considered satisfactory, and this comparison serves as a test for the validity of the theoretical considerations set out in this paper.

For the calculation of densities for emulsion X from equation (16) in table 4, the relevant factors were taken from tables 1 and 3. In table 5, similar calculations for film Y using equations (15) and (8) are shown, the total number of grains per sq. cm. (N) having been calculated from the measured total of AgBr per sq. cm. (table 1) and the value for the average mass per grain (γ in table 3). The general agreement for various wave-lengths is shown in figure 3, and only a limited number of values is given in tables 4 and 5.

Table 4

Film X; $0{\cdot}4343\alpha\bar{a}gk = 4{\cdot}3 \times 10^{-14}$

kv.	Filter	λ effective x.u.	B	Exposure r	d meas.	d calc.
400	2 mm. Sn.	60·5	$3{\cdot}8 \times 10^{11}$	28·0	0·428	0·458
400	Inherent	74	$5{\cdot}8 \times 10^{11}$	27·4	0·683	0·66
190	Thoraeus	119	$1{\cdot}7 \times 10^{12}$	4·5	0·335	0·33
190	1 mm. Al	208	$5{\cdot}2 \times 10^{12}$	0·74	0·184	0·166

Table 5

Film Y; $N = 1{\cdot}37 \times 10^{8}$

kv.	Filter	λ effective x.u.	Exposure r	d meas.	d calc.
190	Thoraeus	119	0·297	1·08	1·08
400	2 mm. Sn	60·5	2·04	1·35	1·46
475	4·5 mm. Sn	47	2·75	1·41	1·5

§ 4. DISCUSSION OF RESULTS

In disagreement with the assumption of previous workers (Silberstein and Trivelli, 1930 ; Charlesby, 1940), the theoretical findings and experimental results described above indicate that the speed of the emulsion is proportional to the product $\eta \times \gamma$ (equation 7 *a*), a conclusion of some significance and calling for a brief discussion.

In the well known absorption formula

$$Q = Q_0 e^{-\mu d},$$

the number of photons (Q_0) crossing a grain will be proportional to the area, whilst the fraction absorbed depends on the thickness and absorption coefficient only. Thus proportionality between the number of photons absorbed and the area would exist only in emulsions consisting either of opaque grains or of grains of equal thickness. The highest absorption coefficient for silver bromide met with in this investigation is of the order 1000 (for $\lambda = 1 \cdot 5$ A.), and taking the diameter of the fairly large undeveloped grain of emulsion Y as $0{\cdot}7 \times 10^{-4}$ cm. gives an absorption of about 7% of the photons crossing a grain, thus excluding any assumption of opacity. The undeveloped grain is regarded as a minute silver-halide crystal in modern theories of the latent image (Gurney and Mott, 1938 ; Mott and Gurney, 1940 ; Berg, 1943), and there seems to be no reason to assume equal thickness of grains.

L. Silberstein and A. P. H. Trivelli (1930) tested the validity of equation (1 *c*) for a series of grain sizes, using single-layer plates, and claimed good agreement between theory and experiment. Re-examination of their results shows, however, that the spread of the experimental values is too large to allow of a clear decision. A brief discussion of the possible dependence of the quantum efficiency on grain-size is given in a later paragraph.

Integration of equation (4) gives

$$n = N(1 - e^{-\eta\frac{\mu}{\rho}\gamma Q}),$$

where the term in the exponent should be equal to the term in the exponent of equation (1 *c*), i.e., we should have

$$\frac{\eta\mu\gamma}{\rho} = a\beta. \qquad \text{......(17)}$$

A. Charlesby (1940) estimated the term *a* from G. E. Bell's measurements of 1936, and found values ranging from $1{\cdot}6 \times 10^{-10}$ to $4{\cdot}8 \times 10^{-10}$ for films which seem to be comparable to film Y. From the data given in Bell's paper, μ can be estimated to be 200, η of the order 10, $\gamma = 1{\cdot}9 \times 10^{-12}$ (see table 3, film Y), and, therefore, $\frac{\eta\mu\gamma}{\rho} = 6 \times 10^{-10}$, in sufficiently good agreement considering the data available as a basis of estimation.

The average energy required to make one grain developable, estimated from the measured quantum efficiencies and corrected for maximum yield from the increase of density for prolonged development (table 3), is 13,000 ev. for film X and 2100 ev. for film Y, which, even with due regard to the error in the determination of the quantum efficiency, is higher than the values given for exposures with visible light. F. R. Hirsch (1928) found that emulsions exposed to X rays of 11 A. (1100 ev.) show the toe, characteristic of exposures to visible light, at the beginning of the density-exposure curve. Taking this as an indication that the quantum efficiency is below unity for this wave-length and the emulsion used (analysis of Hirsch's published graph indicates a quantum efficiency of approximately one-fourth), the same order for the average energy per grain is obtained. Mott and Gurney (1940) estimate that on the average about 50 quanta of visible light are needed to make one grain developable, an average absorbed energy of about 200 ev. per grain. S. E. Sheppard (1931), however, comes to the conclusion that on the average from a few hundred up to one thousand quanta are needed for visible light. The efficiency of the photographic action of x-ray exposures might, however, on account of the almost instantaneous release of a great number of electrons, be nearer to the efficiency of visible light of high intensity, where, as the breakdown of the reciprocity law shows, a higher average energy per grain is required. The theoretical aspect is further complicated by the difference in the energies necessary (film X with a small grain needing more than film Y with a larger grain), which, if not due to the preparation of the films, might be explained on Mott and Gurney's theory by a higher rate of recombination due to higher concentration of electrons in a smaller grain. An interesting aspect of the sensitivity of large grains is presented by W. F. Berg (1943) in a discussion of the electrolytic mechanism of development: "A large grain containing a single sensitivity speck will be more sensitive, not only because of its larger target area to light, but also because it contains a larger total number of interstitial ions to begin with".

Measured values for the exposure in roentgens necessary to produce a density of 0·5, plotted against half-value layer in Cu published by G. E. Bell (1936) and E. E. Smith (1943), were compared with values calculated from equation (16) and show reasonable agreement for harder radiations. The minimum at about 0·25 mm. Cu can be explained by the influence of quanta of lower energy than the K-absorption edge of silver which would have been present in the heterogeneous beams used, and tend to raise the amount of x rays needed before the effective wave-length coincides with the absorption edge.

The theoretical considerations presented in this paper should in principle apply to exposures to visible light. A different mathematical form would have to be chosen to account for the high scattering of light in the emulsion, absorption due to dyes and variation in the quantum efficiency for different grains. A rough estimate of the value of the intrinsic sensitivity (β), using equation (17), can, however, be gained in the following way. J. Eggert and W. Noddack (1923) measured the absorption of light in the silver bromide of photographic emulsions and found a total absorption of 10 to 20%. Assuming a total of 4×10^{-4} gm. AgBr per sq. cm. gives a thickness of $0{\cdot}6 \times 10^{-4}$ cm., which absorbs, say, 15% of the incident light; an absorption coefficient of $2{\cdot}5 \times 10^{3}$ cm.$^{-1}$ is deduced. This high value for the absorption coefficient is explained by the fact that the method of calculation automatically takes the multiple scattering in the emulsion into account. Taking the diameter of the undeveloped grain as 10^{-5} cm. and the quantum efficiency (Sheppard, 1931) as 3×10^{-3} gives $\beta = l\eta\mu = 7{\cdot}5 \times 10^{-5}$. Various workers have arrived at values of 10^{-3} to 10^{-5} for the intrinsic sensitivity.

ACKNOWLEDGMENTS

The author wishes to thank Dr. L. H. Clark and Mr. D. E. A. Jones, B.Sc., for their interest and for their help in the preparation of this paper, and the Medical Officer of Health, Public Health Department, London County Council, for opportunities afforded.

REFERENCES

Arens, H., Eggert, J. and Heisenberg, E., 1931. *Z. wiss. Phot.* **28,** 356.
Bell, G. E., 1936. *Brit. J. Radiol.* **9,** 578.
Berg, W. F., 1943. *Trans. Faraday Soc.* **39,** 115.
Berthold, R., 1925. *Ann. Phys., Lpz.*, 4, **76,** 409.
Blau, M. and Altenburger, K., 1922. *Z. Phys.* **12,** 315.
Charlesby, A., 1940. *Proc. Phys. Soc.* **52,** 657.
Compton, A. H. and Allison, S. K., 1935. *X-rays in Theory and Experiment.* (London: Macmillan).
Eggert, J. and Noddack, W., 1923. *Z. Phys.* **20,** 299.
Glocker, R. and Traub, W., 1921. *Phys. Z.* **22,** 345.
Gurney, R. W. and Mott, N. F., 1938. *Proc. Roy. Soc.* A, **164,** 151.
Hirsch, F. R., 1928. *J. Opt. Soc. Amer.* **28,** 463.
Jones, M. T., 1936. *Phys. Rev.* **50,** 110.
Mott, N. F. and Gurney, R. W., 1940. *Electronic Processes in Ionic Crystals.* (London: Clarendon Press.).
Nutting, P. G., 1913. *Phil. Mag.* **26,** 423.
Sheppard, S. E., 1931. *Photogr. J.* **71,** 313.
Silberstein, L., 1922. *Phil. Mag.* (6), **44,** 257 and 962.
Silberstein, L. and Trivelli, A. P. H., 1930. *Phil. Mag.* **9,** 787.
Smith, E. E., 1943. *Photogr. J.* **83,** 363.
Victoreen, J. A., 1943. *J. Appl. Phys.* **14,** 95.

A METHOD FOR OBTAINING SMALL MECHANICAL VIBRATIONS OF KNOWN AMPLITUDE

By D. H. SMITH,
Woolwich Polytechnic

MS. received 4 *July* 1945

ABSTRACT. The theory of a method for measuring small amplitudes of vibration, due originally to Thomas and Warren, is developed, and experiments confirming it are described. The possibility of using the method for producing a standard source of sound is discussed.

§ 1. INTRODUCTION

The amplitude of a vibrating body such as a diaphragm may be measured either by means of a device which depends upon contact with the body or with an optical arrangement involving the use of interference fringes. The best-known amplitude meters of the first class are the Bragg amplitude meter (Bragg, 1919) and the optical lever used by Kennelly (1923).

Neither of these is very accurate. The first, which depends on the cessation of chattering between the diaphragm and a light spring-loaded hammer, is cumbersome, and cannot well be mounted behind the diaphragm without interfering with the driving arrangements. Kennelly's optical lever is fragile, and needs a powerful arc as a source of light, while its accuracy is low for amplitudes less than about 10 microns. In particular, its usefulness diminishes towards higher frequencies, where a large particle velocity, and, therefore, a high intensity of sound, are obtained with a small amplitude.

Interference fringes were used by Webster (1919), who photographed the fringes produced when a mirror attached to the vibrating body was made one of the mirrors of a Michelson interferometer. A much simpler method was devised by Thomas and Warren (1928), who have described the appearances presented by thin film interference fringes of the Newtonian kind, formed between a reflecting surface at rest, and one set in vibration with a small amplitude. Their description is, however, not quite accurate in several particulars, and the qualitative explanation of the phenomena which they gave appears to have led them to erroneous quantitative conclusions.

In the present paper, the theory of the fringes observed by Thomas and Warren is examined in detail, and a source of sound is described whose amplitude can be accurately measured by their use.

§ 2. THE FORMATION OF INTERFERENCE FRINGES BY REFLECTION FROM THE SURFACES OF A THIN FILM, ONE OF WHICH IS VIBRATING.

Let AB and CD be two close reflecting surfaces of any form, not in contact at any point, and consider the interference of the monochromatic beams OS and OPQS reflected from these surfaces. Let the vibrations in the reflected beams be $y_1 = a_1 \cos \omega t$ and $y_2 = a_2 \cos (\omega t - \delta)$, where a_1 and a_2 are the amplitudes and δ is the phase difference. The resultant is, in complex form,

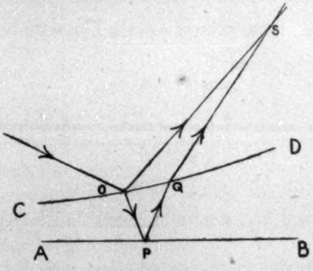


Figure 1.

$$y_1 + y_2 = \exp. j\omega t\,[a_1 + a_2 \exp. -j\delta].$$

The intensity, obtained by multiplying by the conjugate complex quantity, is

$$\begin{aligned} I &= (a_1 + a_2 \exp. -j\delta)(a_1 + a_2 \exp. j\delta) \\ &= a_1{}^2 + a_2{}^2 + a_1 a_2 (\exp. -j\delta + \exp. j\delta) \\ &= a_1{}^2 + a_2{}^2 + 2a_1 a_2 \cos \delta \qquad \ldots\ldots(1) \end{aligned}$$

Let e_0 be the distance between the surfaces at the point O, when they are at rest. Then with sensibly normal incidence

$$\delta = \frac{4\pi e_0}{\lambda},$$

where λ is the wave-length. Let the lower plate vibrate in a sense normal to the path of the light, its motion being $\xi = \xi_0 \cos pt$. The separation of the surfaces at O becomes

$$e_0 + \xi \quad \text{and} \quad \delta = \frac{4\pi}{\lambda}(e_0 + \xi_0 \cos pt).$$

Let

$$\frac{4\pi\xi_0}{\lambda} = \epsilon.$$

Then the intensity at O is

$$\bar{I} = a_1^2 + a_2^2 + 2a_1a_2 \cos(\delta + \epsilon \cos pt).$$

The mean intensity at this point over a period is

$$\begin{aligned} \bar{I}_0 &= \frac{1}{T}\int_0^T I_0\,dt \\ &= \frac{1}{T}\int_0^T [(a_1^2 + a_2^2) + 2a_1a_2 \cos(\delta + \epsilon \cos pt)]\,dt \\ &= (a_1^2 + a_2^2) + \frac{2a_1a_2}{T}\int_0^T \cos(\delta + \epsilon \cos pt)dt \qquad \ldots\ldots(2) \end{aligned}$$

The integral in (2) is

$$I_1 = \int_0^T \cos(\delta + \epsilon \cos pt)\,dt = \cos\delta \int_0^T \cos(\epsilon \cos pt)\,dt - \sin\delta \int_0^T \sin(\epsilon \sin pt)\,dt \qquad \ldots\ldots(3)$$

The second integral in equation (3) is zero, since by writing $\tau = t - T/2$ we see that $\int_{T/2}^{T} = -\int_0^{T/2}$. If we write $pt = \phi$, the first term becomes

$$\begin{aligned} \cos\delta \int_0^T \cos(\epsilon \cos pt)\,dt &= \frac{\cos\delta}{p}\int_0^{2\pi} \cos(\epsilon\cos\phi)\,d\phi \\ &= \frac{\cos\delta}{p}\int_0^{2\pi}\left\{1 - \frac{\epsilon^2\cos^2\phi}{2\,!} + \frac{\epsilon^4\cos^4\phi}{4\,!} - \frac{\epsilon^6\cos^6\phi}{6\,!} + \ldots\right\}d\phi. \end{aligned}$$

Using the well-known reduction formula

$$\int_0^{2\pi} \cos^{2n}\phi\,d\phi = \frac{1\,.\,3\,.\,5\,.\,.\,.\,(2n-1)}{2\,.\,4\,.\,6\,.\,.\,.\,2n}\,.\,2\pi,$$

this becomes

$$\begin{aligned} &\frac{2\pi\cos\delta}{p}\left\{1 - \frac{\epsilon^2}{2^2} + \frac{\epsilon^4}{(2\,.\,4)^2} - \frac{\epsilon^6}{(2\,.\,4\,.\,6)^2} + \ldots\right\} \\ &= \frac{2\pi\cos\delta}{p}\,.\,J_0(\epsilon) = T\,.\,J_0(\epsilon)\cos\delta, \end{aligned}$$

where J_0 is the Bessel function of zero order. Substituting this result in equation (2), we have finally

$$\begin{aligned} I_0 &= (a_1^2 + a_2^2) + 2a_1a_2\,.\,J_0(\epsilon)\cos\delta \qquad \ldots\ldots(4) \\ &= (a_1^2 + a_2^2) + 2a_1a_2 J_0\left(\frac{4\pi\xi_0}{\lambda}\right).\cos\left(\frac{4\pi e_0}{\lambda}\right). \qquad \ldots\ldots(5) \end{aligned}$$

With a given value of ξ_0, which is the same at all points, the intensity varies across the field of view with e_0 and a system of thin-film interference fringes will be formed, whose configuration will depend on the shape of the reflecting surfaces. In the particular case in which we shall be interested, that of fringes formed by reflection from the surfaces of a lens and a flat plate, the usual system of Newton's rings will appear. The contrast between the bright and dark fringes—that is, the visibility—will depend on the value of $J_0\left(\frac{4\pi\xi_0}{\lambda}\right)$. Thus, when $\xi_0=0$, so that $J_0\left(\frac{4\pi\xi_0}{\lambda}\right)=1$, the fringes will vary in intensity between $(a_1-a_2)^2$ in the dark fringes and $(a_1+a_2)^2$ in the bright fringes. But for those values of ξ_0 which make $J_0\left(\frac{4\pi\xi_0}{\lambda}\right)=0$, the intensity will be uniform and equal to $a_1^2+a_2^2$ over the whole field, and for these critical values of ξ_0 the fringes will disappear.

The first few values of ϵ/π and, therefore, of $4\xi_0/\lambda$, for which $J_0(\epsilon)=0$, are shown in the first column of table 1, while the second column shows the corresponding critical amplitude as a fraction of the wave-length of the light used.

Table 1

Critical values of the amplitude

$\frac{\epsilon}{\pi}=\frac{4\xi_0}{\lambda}$	$\frac{\xi_0}{\lambda}$	$\frac{\epsilon}{\pi}=\frac{4\xi_0}{\lambda}$	$\frac{\xi_0}{\lambda}$
0·7655	0·1914	5·7522	1·4380
1·7571	0·4393	6·7519	1·6880
2·7546	0·6886	7·7516	1·9379
3·7534	0·9383	8·7514	2·1879
4·7527	1·1882	9·7513	2·4378

In passing through the value zero, $J_0(\epsilon)$ changes sign, and so, therefore, does the second term in equation (4), so that as the amplitude is gradually increased through a critical value, the fringes disappear and then reappear with the dark and bright bands interchanged in position.

These conclusions are entirely borne out by experiment. Thus, if a well-defined set of fringes is formed with the lower plate at rest, and it is then set in vibration with a very small amplitude, much less than the first critical amplitude, the fringes diminish slightly in visibility. If the amplitude is gradually increased, the contrast diminishes still further, and when the first critical amplitude is reached, the fringes disappear and the field becomes uniformly illuminated. A further increase of amplitude causes the fringe system to reappear with bright and dark bands interchanged. This is repeated as each critical amplitude is reached. The change in the position of the fringes in passing through a critical amplitude is a valuable feature of the effect, as it enables the amplitude to be set for uniform illumination with a precision approaching that obtainable with a good photometer.

The figures in table 1 show that $J_0(\epsilon)$ vanishes for values of ϵ/π, which differ by very nearly unity, and the difference approaches unity closely as ϵ/π increases.

Therefore the successive disappearances of the fringes occur with amplitudes which differ by nearly $\lambda/4$, but the critical amplitudes themselves are not exactly odd multiples of $\lambda/8$, as stated by Thomas and Warren. The theory shows that with high amplitudes the fringes become less distinct, and it becomes slightly more difficult to adjust the amplitude to a critical value. In practice, a_1 and a_2 are very nearly equal, and if they are exactly equal, equation (4) becomes

$$\bar{I}_0^2 = 2a^2[1 + J_0(\epsilon)\cos\delta].$$

For a given value of ξ_0, the visibility of the fringes is then

$$V = \frac{\bar{I}_{\text{max.}} - \bar{I}_{\text{min.}}}{\bar{I}_{\text{max.}} + \bar{I}_{\text{min.}}} = J_0(\epsilon).$$

Negative values of V correspond, of course, to the interchange of the bright and dark fringes on passing through a critical amplitude. The visibility, starting at unity when ξ_0 is zero, reaches maximum values of about 0·4, 0·3, 0·25 and 0·21 between the first and second, second and third, third and fourth, fourth and fifth critical amplitudes, and thereafter decreases much more slowly, so that between the twentieth and twenty-first disappearances it is still 0·1, and between the thirtieth and the thirty-first 0·078. Thus the fringes are still clearly visible even in the neighbourhood of the thirtieth disappearance, and, as described later, the amplitude can still be set accurately to the critical value.

§ 3. EXPERIMENTAL

The conclusions reached in § 2 have been verified with the apparatus shown (figure 2). The vibrator is a flat aluminium-alloy diaphragm (A), 0·025 in. thick and 4·75 in. in diameter, which is tightly clamped to the heavy brass front plate of the instrument by means of a steel ring. Several layers of cartridge paper smeared with hard vacuum grease are interposed between the diaphragm

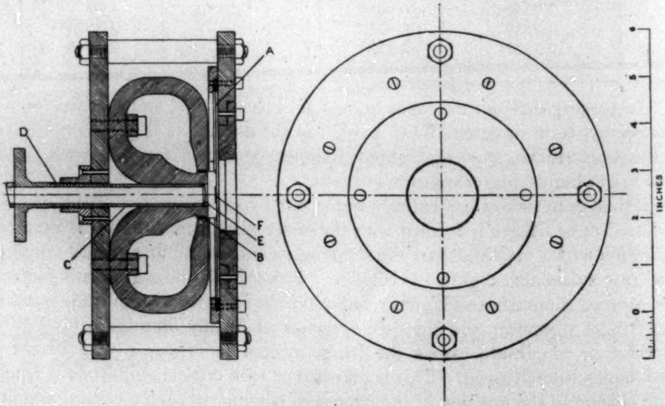


Figure 2.

and the ring, while a single layer of paper is placed between the diaphragm and the front plate ; this is done because in one application of the instrument it is desired to have a gas-tight joint here. The free portion of the diaphragm is 4·5 in. in diameter, and it is separated from the front plate by a distance of 0·005 in. The diaphragm, after being cut from the sheet, was flattened by clamping it between flat steel plates in a large screw press and heating it to about 300° c. The diaphragm is maintained in vibration by means of a "moving coil" drive. The coil (B) consists of two layers, each of 18 turns, of 38 S.W.G. silk-covered copper wire wound on a stiff paper former one inch in diameter, which is cemented to the diaphragm. It moves in the annular gap of a large claw-shaped permanent magnet, of the kind formerly used in moving-coil loud-speakers. The depth of the gap is 0·25 in., while the axial length of the coil is 0·18 in., and the coil is wound in such a position on the former that it is symmetrically situated in the gap. It may be assumed therefore that, when vibrating with small amplitude, the coil remains in a uniform magnetic field, so that a sinusoidal current will produce strictly sinusoidal vibrations. The magnet itself is clamped to a heavy brass back-plate, and the front plate bearing the diaphragm is carried on four stout pillars rising from the back plate. A hole 0·5 in. in diameter is bored through the central soft-iron pole-piece of the magnet, and a close-fitting brass tube (C), whose internal diameter is 0·375 in., can be advanced through it by means of a screw (D) cut on the outside of the tube and working in a nut which is bolted to the back plate. At the end of the tube is mounted a plano-convex lens (E), the radius of curvature of whose front surface is 200 cm. The lens is mounted on a brass plate which can be tilted by means of three small spring-loaded screws; for the sake of clarity, this arrangement is not shown in the diagram. At the centre of the diaphragm, opposite the lens, is cemented a circular microscope cover-slip (F), whose back face is lightly ground and coated with optical black to abolish unwanted reflections. With the lens close to the glass plate, but not quite touching it, Newton's rings are formed by reflection at their surfaces when monochromatic light is projected down the tube. The screw has fifty threads per inch, so that the distance between the lens and the plate can be delicately adjusted. The tilting screws are manipulated with a long screw-driver passed up the tube, and with a little practice it is easy to bring the centre of the rings to the middle of the field of view, while by turning the tube, the separation of the lens and the plate is adjusted to be about 10 microns. The fringes are viewed through the tube with a telemicroscope.

An aluminium cover, not shown in the diagram, encloses the rear part of the apparatus. As the whole instrument is massively constructed and the moving parts are accurately fitted, it is scarcely affected by shock, and the only effect of tapping the case is to produce a slight quiver of the fringes, which is due to free vibration of the diaphragm.

Observations were made at a number of frequencies of the R.M.S. currents required to produce disappearance of the fringes, i.e. at the critical amplitudes. The current was measured with a Unipivot galvanometer with a square-law scale and a series of vacuum thermojunctions whose ranges overlapped, so that the current readings at each critical amplitude could be made with nearly the same accuracy. The thermojunctions were calibrated with direct current, using a direct-reading potentiometer and standard resistances.

The observations were made with sodium light, whose lack of monochromacity introduces no appreciable error into the results.

§4. RESULTS AND DISCUSSION

It was found that the amplitude could be set to a critical value with an accuracy of about one-half per cent, as judged by the agreement between the values of the current with successive settings. At the higher amplitudes, the falling-off in visibility is more than offset by the fact that a given change in amplitude is brought about by a smaller fractional change of current than at lower amplitudes.

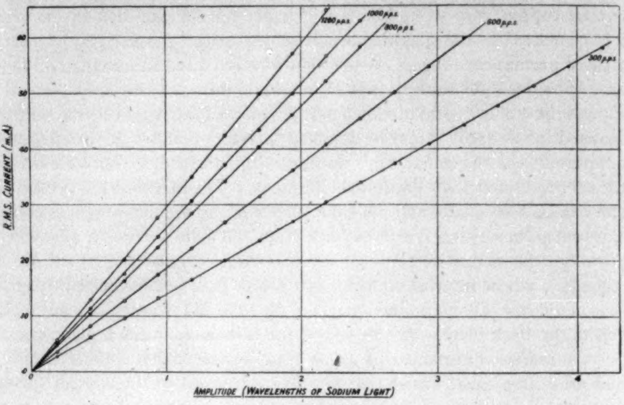


Figure 3.

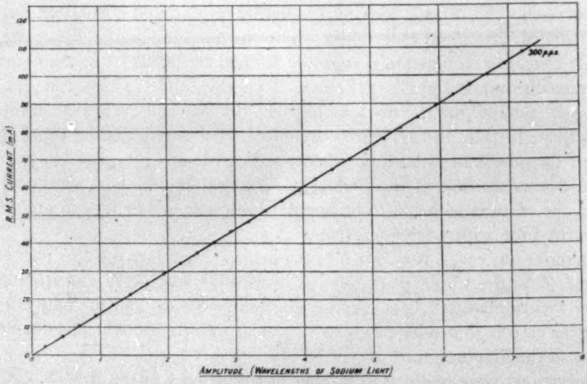


Figure 4.

The results at five frequencies are shown graphically in figure 3, where the current for which the fringes disappear is plotted against the corresponding critical amplitude given by the calculation of § 2. Each point represents the mean of four observations. In each case the curve relating current and amplitude is accurately a straight line passing through the origin, thus confirming the theory.*

It is evident, from the accuracy with which the points lie on a straight line, that when the best straight line through them has been found, the amplitude can be set to any assigned value within the range by adjusting the current, with an error not greater than 1 in 500. In the case of the observations at 300 c./sec., the equation of the straight line through the origin and the centroid of the points is

$$\mathscr{I} = 13{\cdot}900(\xi/\lambda),$$

where $\mathscr{I}$ is the current, and the maximum deviation (on the current axis, since the amplitude values are pre-assigned) which occurs at the seventeenth disappearance, with a current of 58 ma., is 0·2 ma.

In figure 4 is shown the result of observations extending to the thirtieth disappearance ($\xi_0 = 4{\cdot}38\,\mu$). It would appear justifiable to extrapolate over double this range. The amplitude would then be about 8 μ, and there is no reason to suppose that the law connecting current and amplitude would cease to be a linear one.

The instrument shown in figure 2 has been designed primarily for use in a new tube method for measuring the velocity of sound, but with slight modification it constitutes a standard source of sound, whose amplitude can be set to have a known absolute value with high accuracy. It has been described here in the form with which the observations for figures 3 and 4 were made, but a more compact pattern has been built, in which an axially magnetized cylindrical magnet is used. In order to make it suitable for use as a standard source, it is fitted with a thinner front plate, and the diaphragm carries a composite piston of balsa wood and aluminium, whose front face is coplanar with the front plate. When it is mounted in an extensive baffle in an acoustic chamber, it produces a sound field whose distribution can be calculated. The R.M.S. displacement velocity at a point on the axis l cm. in front of such a piston, of radius r, is given by

$$\hat{\dot{\xi}} = \frac{\dot{\xi}_0}{\sqrt{2}}\left\{\left(\frac{l}{\alpha}\right)^2 + 1 - \frac{2l}{\alpha}\cos\frac{2\pi}{\lambda_1}\left(\alpha - l\right)\right\}^{\frac{1}{2}},$$

where $\dot{\xi}_0$ is the peak displacement velocity of the piston, λ_1 is the wave-length of the sound, and $\alpha = \sqrt{(l^2 + r^2)}$ (West, 1932). Table 2 shows the displacement velocity on the axis of the piston, of radius 2 cm., at a distance of 1 metre, together with the approximate intensity in db. above 10^{-16} w. per cm.², when the amplitude of the piston is 4 μ.† This corresponds to about the thirtieth critical amplitude with mercury green light ($\lambda = 5461 \times 10^{-8}$ cm.) The R.M.S. displacement velocity corresponding to the reference level is about 5×10^{-6} cm./sec. at 1000 c./sec. As ordinary speech corresponds to an intensity level of about 40 db. at 1 metre, the loudness is adequate for acoustic measurements without extrapolating from calibration curves such as figure 4, except at the lowest frequencies.

---

* The apparent discrepancy in the slope of the line per 1000 c./sec. is due to the fact that some adjustments were made to the apparatus between this set and the other sets of readings.

† A sound has an equivalent loudness of *n* B.S. phons when it has the same loudness (judged under specified conditions) as a standard tone in the form of a plane sinusoidal wave at 1000 c./sec., whose intensity level is *n* db. above a reference level corresponding to an R.M.S. sound pressure of 0·0002 dynes per cm². This corresponds very closely, in air, to an average power of 10^{-16} w. per cm², and the intensities above have been referred to this level, to obtain a rough comparison with speech.

Table 2

Intensity of sound from the diaphragm at a distance of 1 metre

f (c./sec.)	$\hat{\dot{\xi}} \times 10^4$ (cm./sec.)	Intensity (db. above 10^{-16} w. per cm^2)
100	0·654	22·5
500	16·37	50·3
1000	65·4	62·5
2000	262·0	74·4

Experiments are in progress on the use of the source for calibrating microphones, and for other acoustic measurements.

Drysdale (1939) emphasized the need for a standard source of sound, and suggested means by which a spherical source might be constructed, but he pointed out that the chief difficulty would lie in the measurement of the amplitude, which would need to be of the order of 0·2 μ. This is well within the scope of the present method.

ACKNOWLEDGMENTS

I am indebted to the Governing Body and the Principal of the Woolwich Polytechnic for providing facilities which enabled this work to be done. I have also to thank High Duty Alloys, Ltd., for presenting a quantity of aluminium-alloy sheet for diaphragms when war-time restrictions made it difficult to obtain this material, and to Darwins, Ltd., for providing some magnets in similar circumstances.

REFERENCES

Bragg, W. H., 1919. *Engineering*, **107**, 777.
Drysdale, C. V., 1939. *Proc. Phys. Soc.* **51**, 359.
Kennelly, A. E., 1923. *Electrical Vibration Instruments* (New York, ed. by D. C. Jackson and E. R. Reddick).
Thomas, H. A. and Warren, G. W., 1928. *Phil. Mag.* **5**, 1125.
Webster, A. G., 1919. *Proc. Nat. Acad. Sci., Wash.*, **5**, 179.
West, W., 1932. *Acoustical Engineering*, 69.

THEORETICAL INVESTIGATION ON TELEPHOTO LENSES

By T. SMITH, F.R.S.,
Teddington

Paper read to the Optical Group 30 *April* 1945

ABSTRACT. The theory of thin lenses is employed to investigate possible forms of telephoto lenses consisting of two widely separated thin components. From a consideration of the shapes of the unit surfaces it is shown that permissible forms have a limited range on either side of those in which both components are aplanatically corrected ; on the one side the lenses tend to present their more convex aspects to one another, on the other side their more concave aspects. On taking other conditions into account it is found that heavy figuring is required except in two cases. In one of these the components are nearly aplanatically corrected, and in the other they are strongly concave towards one another. The former construction has the advantage in offering less strongly curved surfaces. It is shown that a design of this class, consisting of two ordinary cemented doublets each made from the same kinds of glass, has stable characteristics.

The formulae used in the investigation are given in an appendix.

§ 1. INTRODUCTION

During the period of the war the focal lengths of photographic lenses used for reconnaisance have been progressively increased, and to meet this requirement large-scale copies of telephoto lenses designed to meet somewhat different conditions have been made and used extensively. In view of the importance of securing the best possible photographs it nevertheless appeared desirable to carry out a theoretical investigation on telephoto lenses to see whether improvements were likely to be attainable, bearing in mind the work for which these lenses are required. Briefly, the requirements are the attainment of the finest possible definition over a field of angular diameter about 25°, and the comparative unimportance of distortion. For some other uses to which telephoto lens are put, the elimination or reduction of distortion is important, and it may be that to attain this, something which is important in a lens for aerial photography is sacrificed in the normal type of telephoto lens.

It is of great importance that the photographs obtained with these lenses should suffer as little as possible from loss of contrast. With designs otherwise equally good, this is most likely to be secured by a lens with the smallest possible number of glass-air surfaces. It is therefore appropriate to develop the theory in the expectation that neighbouring lens surfaces will be cemented together if satisfactory definition with a suitable aperture can be obtained with this construction. The initial assumption is therefore made that the telephoto combination will consist of two lenses, the leading one of positive power and the other of negative power, each preferably composed of cemented components. The work proceeds on the basis that each of these lenses may, in the first place, be treated as of negligible thickness. As the combination must be chromatically corrected for both position of focal plane and size of image, it is essential that each cemented lens should be corrected independently for colour. As the lenses are thin this involves only one condition for each lens.

Correction of the combination is secured by imposing suitable conditions on each of the two components. The problem in fact is reduced to the design of two thin lenses. To follow the theory it is therefore important to bear in mind certain properties of thin lenses. For convenience the power of the thin lens is taken as unity.

§ 2. PROPERTIES OF THIN LENSES

Since this lens is to be achromatic at least two glasses with different ν values must be used. The glass with the greater ν is used for an element of positive power, that with the smaller ν for an element of negative power. The algebraic sum of these powers is unity : let N be the sum of their absolute values. Then the lens will be achromatic if

$$N=\frac{\nu_1+\nu_2}{\nu_1-\nu_2}.$$

So far as this condition is concerned it is immaterial whether each kind of glass is represented by a single element or by several elements, whether components are cemented or successive surfaces have different curvatures where they touch on the axis ; the order in which the elements are arranged and their shapes are also of no significance. N can therefore be regarded as a current variable when monochromatic aberrations are considered ; it will only be necessary at some stage to refer to a catalogue of optical glasses to verify that types giving the value of N desired are available. Negative as well as positive values of N may be admitted, the negative sign indicating that the glass which it was expected would have the smaller value of ν must, in fact, have the greater value. No loss of generality is now involved in supposing that with the definition of N already given ν_1 is the constringence of the glass of lower refractive index, conventionally referred to in what follows as crown glass, and ν_2 that of the glass of higher refractive index, to be referred to as the flint glass. The particular meaning assigned to crown and flint must be kept in mind. For instance, the "crown" may prove to be a glass described in the glass catalogue as a light flint, and the "flint" as a dense barium crown. In such a case the value of N would be negative.

The first-order monochromatic aberrations are usually given as

Spherical Abberration,
Coma,
Astigmatism and Curvature,
Distortion.

Curvature belongs to a different group of aberrations from the other four, and is measured by the Petzval sum for the lens. This is an invariant, independent of the order of the components and their shapes. In a thin lens of two glasses it is a linear function of N. For a thin lens the aperture stop is presumed to be placed in contact with the lens itself. Taking the aberrations in the reverse order to that given above (for thin lenses at least this reversed order is the simpler of the two), the distortion is identically zero, and the astigmatism is invariable and has the same value for all lenses irrespective of their construction. The astigmatism and curvature together determine the shapes of the focal surfaces for sagittal and tangential lines; they are represented in section in the centre of figure 1 by the circular and oval curves. In any satisfactory instrument for photographic use the lenses must be combined in such a way that the focal surfaces, instead of having these most unsuitable shapes, are substantially coincident over he angular field to be covered.

The two remaining aberrations are under control in different ways. The essential requirement for the elimination of coma is to secure the correct shape for the unit surfaces of the lens, and for the removal of spherical aberration to ensure equality in focal length for rays through different parts of the lens aperture. Some of the means at our disposal for satisfying these conditions in cemented lens combinations are illustrated in figures 1 and 2.

The lenses shown in figure 1 are externally symmetrical—all the external surfaces have the same curvature. Starting from the doublet at the top of the diagram, with the crown lens leading, the parabolic track on the left leads through forms in which the crown lens is divided, the value of the power of the crown glass transferred from the left to the right side of the flint lens being proportional to the distance descended in the diagram. At the extreme left is shown the symmetrical form in which the two crown components are of equal power. Thereafter the power of the crown component on the left diminishes until at the bottom of the diagram we reach the doublet form with the flint component leading. Lenses could be constructed on the parabola beyond the doublets, but they are less desirable forms, as the curvatures of the internal surfaces are greater than in other forms. The lenses on the other parabola indicate a second series in which the flint lens is divided instead of the crown. Other forms can be constructed in which both flint and crown are divided, so that there are at least four components, but these will not be used in manufacture if simpler forms prove adequate. The aberrational properties of all these externally symmetrical lenses

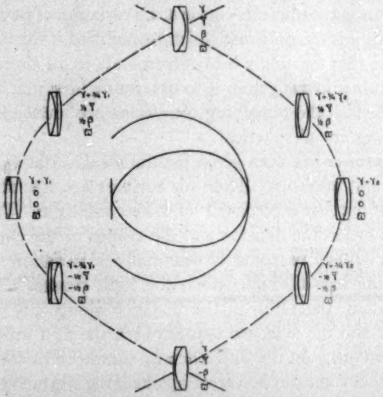

γ
$\bar{\gamma}$
β
ϖ
$\gamma+\frac{1}{4}\gamma_1$
$\frac{1}{2}\bar{\gamma}$
$\frac{1}{2}\beta$
ϖ
$\gamma+\frac{1}{4}\gamma_2$
$\frac{1}{2}\bar{\gamma}$
$\frac{1}{2}\beta$
ϖ
$\gamma+\gamma_1$
0
0
ϖ
$\gamma+\gamma_2$
0
0
ϖ
$\gamma+\frac{3}{4}\gamma_1$
$-\frac{1}{2}\bar{\gamma}$
$-\frac{1}{2}\beta$
ϖ
$\gamma+\frac{3}{4}\gamma_2$
$-\frac{1}{2}\bar{\gamma}$
$-\frac{1}{2}\beta$
ϖ
γ
$-\bar{\gamma}$
$-\beta$
ϖ

Figure 1. Aberrational coefficients of representative lenses having the same focal surfaces.

and of others which can be derived from them involve six coefficients which are functions only of N and of the refractive indices of the glasses. Of these coefficients, ϖ, the Petzval sum, enters with equal weight in all forms. The next coefficient, β, has weight proportional to the internal asymmetry of the lens, and has its sign reversed if the lens is reversed. As β denotes the sum of the curvatures of the unit surfaces of the lens, this change of sign on reversal is to be expected, as is also the zero coefficient of β for the completely symmetrical forms. The coefficient $\tilde{\gamma}$ behaves in a similar way to β for reasons which will become clear from consideration of a relation to be given later. The ratio of $\tilde{\gamma}$ to β is, in fact, invariable for glasses of given refractive indices. $\tilde{\gamma}$ and β are quadratic in N and vanish when there is only one kind of glass, i.e. when $N = \pm 1$. γ, γ_1 and γ_2 are cubics in N. The roots of $\gamma_1 = 0$ are $1, -1, -1$, and those of $\gamma_2 = 0$ are $1, 1, -1$; $\gamma = 0$ has only one real root. The weight of γ does not change, but, on proceeding in either direction from the symmetrical forms, $\gamma + \gamma_1$ and $\gamma + \gamma_2$ are diminished by $\theta^2\gamma_1$ and $\theta^2\gamma_2$ as the case may be, where θ is the weight given to β and $\tilde{\gamma}$. All the gamma coefficients concern the variation of power with aperture. Of all six coefficients γ is the only one which is modified if the refracting surfaces are not spherical, so that figuring a lens corresponds to an alteration in γ. The method followed in this investigation is to determine how much figuring of each lens is required when all the remaining conditions are satisfied. If no figuring is needed the surfaces are all spherical.

The significance of ϖ has been given for the focal surfaces, but it is equally important for the unit surfaces. The unit surfaces are free from astigmatism, but the curvature of the first exceeds that of the second by ϖ. This result again is independent of the order, the division and the shapes of the component lenses.

The aberration known as coma is essentially a variation of magnification with the zone of the lens through which the light passes. To maintain this magnification constant the ratio of the distance along any ray between the second unit surface and the image to that between the first unit surface and the object must be constant. As the difference in curvature of the unit surfaces is fixed, it is obvious that coma can be varied by changing the curvature of both unit surfaces together. If, instead of being zero, as in figure 1, the sum of the curvatures of the external surfaces of the lens is made equal to ρ by increasing the curvature of every refracting surface by $\frac{1}{2}\rho$, the increment in the sum of the curvatures of the unit surfaces is $\rho(1+\varpi)$, so that if this sum were initially $\theta\beta$ it would become $\theta\beta + \rho(1+\varpi)$ after bending. Figure 2 shows a number of forms of lens, both doublets and triplets, bent by a number of equal increments, with the corresponding shapes of the unit surfaces in the last line. In the central column the triplets are both symmetrical, and the curvatures of the unit surfaces equal and opposite. The positions that would be occupied by cemented doublets with equal and opposite curvatures for their external surfaces are indicated by the asterisks in the first and fourth rows.

From the explanation just given it can be readily shown that the condition for freedom from coma when the magnification is m is that the sum of the curvatures of the unit surfaces is $M(2+\varpi)$, where

$$M = \frac{1+m}{1-m},$$

i.e. it is

$$\theta\beta + \rho(1+\varpi) = M(2+\varpi).$$

When the coma has been removed in this way, the condition that the lens should be free from spherical aberration is

$$\gamma + \gamma_1 - \theta^2\gamma_1 + 2\rho\theta\tilde{\gamma} + \rho^2(1+2\varpi) = M^2(5+2\varpi),^*$$

the lens being a triplet with the crown component divided in the ratio $1+\theta : 1-\theta$. If the flint is divided instead of the crown, γ_2 must be written in place of γ_1. At the moment it is only necessary to note that there are two degrees of freedom

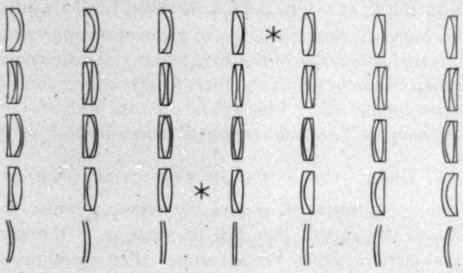


Figure 2. Effect of bending lens on curvatures of unit surfaces.

and that in general two conditions can be satisfied. As one of the equations is quadratic it will sometimes happen that the solutions are imaginary.

Returning now to the list of aberrations, and remembering that curvature of field has to be dealt with separately, there are four conditions to be satisfied by the four degrees of freedom in the two thin lenses. But for the reason already given the condition for freedom from distortion will be disregarded and one degree of freedom retained to be used in whatever manner seems most advantageous. The free variable secured in this way will be denoted by ϕ.

§ 3. THE FUNDAMENTAL STRUCTURE OF THE TELEPHOTO LENS

Before proceeding further it is necessary to consider what the powers of the two component lens groups are to be, and how far they are to be apart. The Petzval sum for the complete lens, which is to vanish, is the sum of those for the components. The ratio of the Petzval sum to the power of a thin lens is only variable to a small extent, so that the powers of the two components will tend to be in a fixed ratio, which is approximately -1. Suppose that the powers of the two are a and $-ma$. If the power of the combination is unity, the separation t satisfies

$$a - ma + tma^2 = 1$$

and the overall length from the first lens to the focal plane is

$$1 + t - ta$$

---

* As the value of the power does not involve any distinction between object and image spaces, this condition should not change when the lens is reversed, i.e. when ρ and M are changed in sign. It follows that $\tilde{\gamma}$, like β, changes in sign, but not in magnitude, when the lens is reversed.

or

$$1-\frac{m}{4}+\frac{1}{m}\left(\frac{1}{a}-1+\frac{m}{2}\right)^2.$$

It follows that, for minimum overall length,

$$a=\frac{2}{2-m}, \quad t=\frac{2-m}{4}.$$

In all telephoto lenses of minimum length the separation is therefore half the focal length of the positive component, and the back focal length is half the focal length of the complete lens. It has been customary to describe a telephoto combination with the focal length double the back focal length as a two-times system. In this sense all telephoto lenses of greatest compactness are two-times systems, and this limits the value of the description. An alternative meaning for a two-times system is one in which the focal length of the combination is twice that of the positive component. This will be the case with $m=1$. The systems to be considered here are two-times telephoto lenses in both senses.

§4. UNIT SURFACES OF THE COMPONENTS

With the two components of numerically equal powers, it is appropriate n the first place to assume that they will be made from the same glasses and, therefore, have numerically equal Petzval sums. For a preliminary exploration it may be assumed that the value of the Petzval sum for a thin lens is 0·75 times the power. One form of construction which satisfies the conditions laid down consists of components aplanatically corrected for the magnification at which they are operating. It is clear that this secures freedom from coma and spherical aberration, and under these conditions the astigmatic contributions, which are proportional to the powers of the thin lenses, are also additive, so that both first-order astigmatism and curvature are eliminated. Since for the positive component $m=0$, and for the negative $m=2$, the corresponding values of M are 1 and -3, and the sums of the curvatures of the unit surfaces of these two components are required to be

$$2\times 1\times(2+\cdot 75)=5\cdot 5$$

and

$$(-2)\times(-3)\times(2+\cdot 75)=16\cdot 5.$$

Their differences are $2\times\cdot 75$ and $-2\times\cdot 75$ respectively, or 1·5 and $-1\cdot 5$, giving for the curvatures of the four unit surfaces in order the values

$$3\cdot 5,\ 2\cdot 0,\ 7\cdot 5,\ 9\cdot 0.$$

When the components are not aplanatic, it may be shown that the sum of the curvatures of the first should be increased by $\frac{1}{2}\phi$ and that of the second by -2ϕ. The curvature of the second unit surface of the complete lens must be unity, since the focal length is to be the same for all rays incident parallel to the axis. Figure 3 shows the dependence of all these curvatures on ϕ.

It is profitable to exhibit the variation of the unit surfaces with ϕ in another way, shown in figure 4. Consider first the case marked $\phi=0$. On the extreme left two rays incident parallel to the axis are represented by full lines. They travel in this direction until they meet the first unit surface of the positive lens. (Owing to the sign of the difference in the curvatures of the unit surfaces, the first unit surface appears on the right of the second, when the light travels from left to right, if, as here, the surfaces are in contact on the axis.) The rays then converge from the points of the second unit surface at the same distance from the axis to the paraxial focus of the first lens. After travelling about half-way to this point, these rays meet the first unit surface of the negative lens, and thence proceed from the corresponding points of the second unit surface to the focus of the

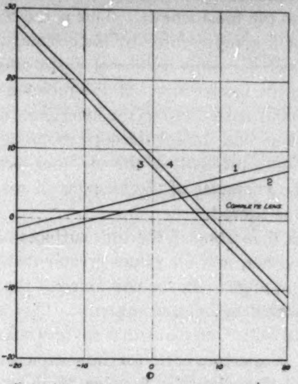


Figure 3. Curvatures of unit surfaces.

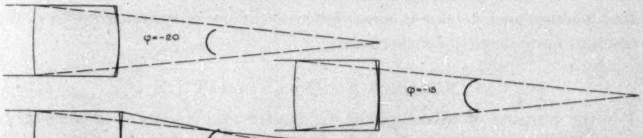


Figure 4. Variation of unit surfaces of component lenses with ϕ.

complete lens. (As the second lens is negative the unit surfaces appear in their natural order.) If now the emergent rays are produced backwards to meet the incident rays, the intersections will be found to lie on the sphere, centred at the focus, of radius equal to the focal length. This is the second unit surface of the complete instrument, and is represented by the circular arc on the extreme left.

In the diagrams drawn for other values of ϕ the course of the rays may be determined by finding the intersections of the incident rays with the first unit surface of the positive lens, and those of the emergent rays with the second unit surface of the negative lens, and joining the corresponding points of the remaining unit surfaces to determine the path of the rays between the two lenses. The figure shows clearly the character of the changes in spherical aberration of the component lenses as ϕ is altered.

For values of ϕ less than about 5 the unit surfaces of the two lenses tend to turn away from one another, and for values greater than 5 towards one another. For extreme values of ϕ, e. g. -20, the curvature of the second lens is too great to allow the rays indicated to be transmitted. This is substantially the case also with $\phi = -15$ and $\phi = 25$. Systems that are to work at a fairly large aperture (the rays are drawn for an aperture of f/5 for the complete lens) must correspond to values of ϕ well within these limits. A value of about 5 would be very satisfactory. (For freedom from distortion the value would be 16 with $\varpi = 0{\cdot}75$.) To find whether such a value is consistent with the use of spherical surfaces alone necessitates more detailed consideration.

§ 5. NUMERICAL INVESTIGATION

For the purpose of investigating the matter numerically it is necessary to take into account the magnitude of the functions of the refractive indices which have already been mentioned. No advantage will be gained at this stage by basing the calculations on the properties of glasses given in manufacturers' catalogues. Instead, two refractive indices are chosen as is thought convenient, and N is varied. The results which follow are for a crown glass with $\omega_1 = 0{\cdot}64$ and a flint with $\omega_2 = 0{\cdot}60$, where ω is the reciprocal of the refractive index. It is much more convenient in dealing with aberrations to take the reciprocal than the refractive index itself.

The calculation of the surface curvatures can be carried out in any of the familiar ways. In the method used at the National Physical Laboratory they are obtained incidentally in the calculation of the six aberrational coefficients. These curvatures for the unbent forms of doublet (with crown glass leading) and for both kinds of symmetrical triplet are shown in figure 5. The lines marked "First" and "Last" represent the curvatures of the external surfaces in all these lenses. That marked D gives the curvature of the cemented surface of the doublet. The diagram shows that it is important to discover a system which gives a fairly small value of N—a value numerically less than 5 is very desirable. The lines marked "$2(\gamma_1)$" and "$3(\gamma_1)$" give the curvatures of the cemented surfaces when the crown lens is divided, and similarly those marked "$2(\gamma_2)$" and "$3(\gamma_2)$" the curvatures when the flint lens is divided. As would be expected, these curvatures are much less than that of the cemented surface of a doublet, particularly when the positive component is divided. It does not necessarily follow that this advantage is retained in the final form of the lens, for the triple lens may need to be bent more than the doublet. Any bending is represented on this diagram by a change in the position of the zero of the curvature scale.

Figure 6 shows how the four coefficients required to determine the properties of cemented doublets depend on the value of N. The changes in γ are much more important than those of the other functions. It is mainly the value of γ which determines the suitability of a combination of glasses for any given purpose. This value of γ is to be compared with the value which is required to correct the system. The equations to be used are, for the first component,

$$\beta+\rho(1+\varpi)=2+\varpi+\tfrac{1}{4}\phi,$$
$$\gamma+2\rho\tilde{\gamma}+\rho^2(1+2\varpi)=5+2\varpi+2\phi,$$

and for the second component

$$\beta+\rho(1+\varpi)=-3(2+\varpi)+\phi,$$
$$\gamma+2\rho\tilde{\gamma}+\rho^2(1+2\varpi)=9(5+2\varpi)+4\phi.$$

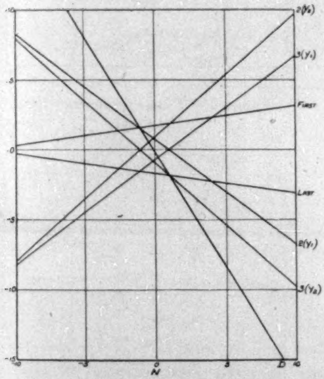


Figure 5. Curvatures of unbent lens surfaces.

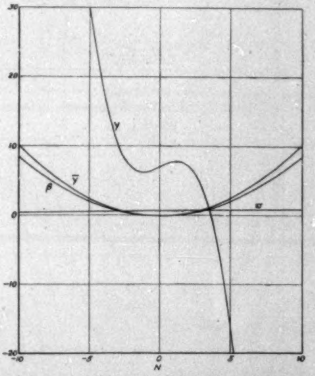


Figure 6. Aberrational coefficients for cemented doublets.

The first equation of each pair is used to determine the value of ρ, and the desirable value of γ is found by substituting for ρ in the second equation. As they stand, the equations apply to components with the crown lens leading ; when the flint lens leads the signs of β and $\tilde{\gamma}$ are to be changed.

Figure 7 shows the curves giving the desirable values of γ for several values of ϕ ; the cubic curve giving the value of γ for these glasses when all the surfaces are spherical is repeated from figure 6. It will be seen that the value of N where the last curve intersects the others extends from about $N=3$ upwards. The minimum corresponds to a value of ϕ of about 8. Since N is positive for all intersections, the glasses to be used are a normal crown and a normal flint. For the minimum value of N a somewhat extreme pair would be required, but good combinations seem possible for values of ϕ between, say, -6 and 24. Figure 8 shows the corresponding curves for the positive component when the flint lens leads. The deductions to be drawn in the two cases are very similar.

The vertical displacement between any ϕ curve and the cubic curve for any value of N indicates the amount of figuring of the lens that would be needed. It is immaterial whether this is done on the first or the last surface. The scale is such that when $N=1$ the distance of the cubic curve above $\phi=0$ represents the figuring that would be needed to make a single lens of index 1·625 aplanatic, and when $N=-1$ that required if the index were 1·66667.

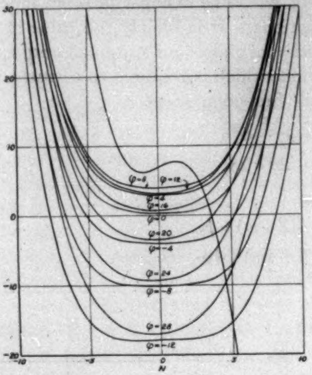


Figure 7. First component: crown leading.

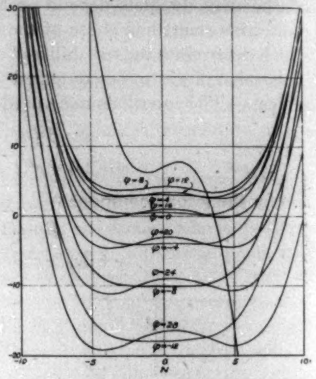


Figure 8. First component: flint leading.

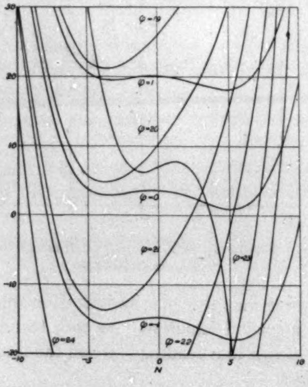


Figure 9. Second component: crown leading.

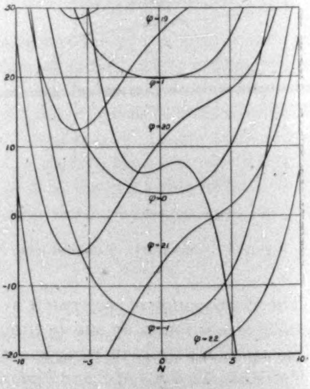


Figure 10. Second component: flint leading.

Figure 9 gives results for the second component with crown lens leading. The curves are here much more widely spaced, and the minimum value of N is negative and probably numerically too large to be attractive. Positive values of ϕ less than 20 that involve no figuring would have to be made from a pair of glasses with the greater dispersion in that of lower refractive index. For negative values of ϕ or positive values greater than 21 a normal pair of glasses would be used. If a negative value of N were accepted as a basis of construction it would probably be desirable to modify the initial assumptions to the extent of using a negative component of somewhat greater power than the positive, and thus enable a very small value for the total Petzval sum to be maintained.

Figure 10 shows the curves for the negative lens when the flint glass leads. The inferences to be drawn are very similar to those from the previous case.

Doublet lenses represent one of the extremes within which it is desirable to work. Symmetrical triplets represent the other extreme. Figure 11 shows in full lines the curves for γ_1 and γ_2 as well as for γ, and in broken lines those for $\gamma+\gamma_1$ and $\gamma+\gamma_2$. It is possible to make from given glasses a cemented lens having properties intermediate between those indicated by the appropriate doublet and triplet curves.*

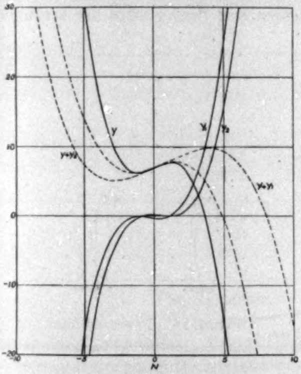


Figure 11. Aberrational coefficients for triple lenses.

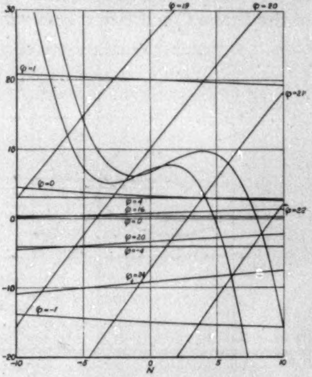


Figure 12. Symmetrical triple lenses.

Figure 12 gives curves for both forms of symmetrical triplets. Those relating to the positive component have the values of ϕ recorded in the middle of the diagram. The curves for the negative component have their ϕ values inserted near the margin of the diagram. These curves are to be considered in connexion with the $\gamma+\gamma_1$ curve copied from figure 11 if the crown lens is split, and with the

---

* Flexibility can often be secured by not approaching either extreme very closely. For example, it has frequently been required at the National Physical Laboratory to construct highly corrected aplanatic objectives, usually two or three in number, for any given specification. If cemented doublets were made (and cementing is practically essential in many instruments for making fine measurements) it would be necessary to obtain glasses having optical properties conforming to a fairly rigid specification; but this is no longer a requirement if glasses well away from the doublet region on the side of the triplet boundary are chosen. The neighbourhood of the doublet boundary is avoided to prevent one part of the divided crown lens from being of an inconvenient shape—it would probably be a deep meniscus for glasses near the doublet locus. Where large-scale production is involved other considerations become important.

$\gamma+\gamma_2$ curve if the flint lens is split. It is not obvious from this diagram that any marked advantages are obtainable by the use of triplet components.

In addition to the values of the six aberrational coefficients, their chromatic variations are important. These are shown in figure 13. In applications an additional factor $1/(\nu_1-\nu_2)$ has to be applied to obtain directly comparable figures for systems made with different glasses. The outstanding points are the relative unimportance of the variations in ϖ, β and $\tilde{\gamma}$, and again the value of glass combinations in which N is numerically small.

From inspection of the curves of figures 7, 8, 9, 10 it appears that telephoto lenses in which the same kinds of glass are used in both components are possible, and would be fairly representative of practicable systems generally. These systems have been determined for all doublet combinations using the glasses $\omega_1=0{\cdot}64$, $\omega_2=0{\cdot}60$, and details are given in the accompanying table. The order of the letters C and F is that in which the crown and flint glasses are arranged in the complete system. The first point to be observed is that the values of ϕ are very restricted. Effectively there are only two choices—either ϕ is nearly zero or it is about 21. As figure 4 indicated, the former value is the more promising. The table below shows that the smaller value of ϕ carries with it a smaller value of N, which is also a favourable indication. From the curvatures of the refracting surfaces of the eight possible systems it will be seen that the one with the least steep surfaces has a small negative value of ϕ and is arranged with the two crown glasses facing outwards. The next best system has a large ϕ, and both crown lenses precede the flint lenses to which they are cemented. A close third is the system with small ϕ and the flint components preceding the crowns. The form of construction which, on this analysis, appears the most favourable is illustrated in figure 14.

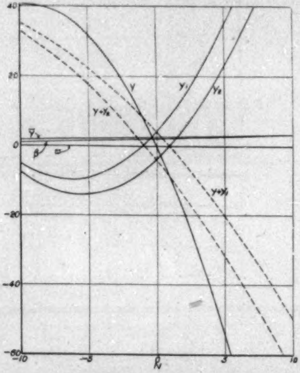


Figure 13. Chromatic variation of aberrations.

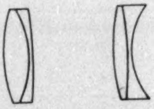


Figure 14. Preferred form of lens.

Having selected a type of lens for fuller examination, it is of interest to know whether the construction will be very sensitive to changes in the refractive index of the glass. To ascertain this it is sufficient to vary one glass only, and the crown

Table

Two-times telephoto lenses of unit focal length
Glasses $\omega_C = 0{\cdot}64$, $\omega_F = 0{\cdot}60$

	Glass order			
	CFCF	CFFC	FCCF	FCFC
N	3·407	3·434	3·580	3·607
ϕ	−0·019	−0·247	−0·083	−0·338
R_1	3·160	3·120	4·317	4·292
R_2	−4·675	−4·762	8·187	8·202
R_3	−1·064	−1·111	0·044	0·012
R_4	3·217	2·257	3·288	2·222
R_5	11·058	−1·394	11·430	−1·688
R_6	7·442	6·488	7·560	6·502
N	3·802	3·749	4·358	4·274
ϕ	21·518	21·053	22·024	21·377
R_1	6·245	6·192	7·955	7·824
R_2	−2·288	−2·250	12·905	12·735
R_3	1·912	1·873	3·485	3·359
R_4	−8·133	− 9·028	−9·472	−10·880
R_5	0·400	−13·151	−0·050	−15·791
R_6	−3·800	− 4·709	−5·000	− 6·415

The *R*s denote the curvatures of the lens surfaces.

is an obvious choice. Figure 15 shows the curvatures of the six surfaces as the reciprocal crown index varies from 0·625 to 0·68 : the corresponding index limits are 1·6 and 1·47. It is clear that the stability of this type of lens is satisfactory.

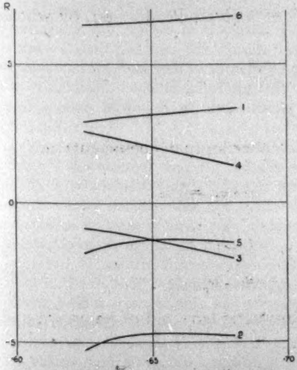


Figure 15. Changes in curvatures of C F F C lens as crown glass is altered.

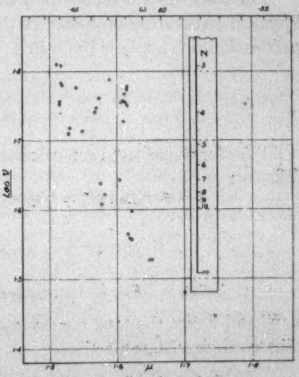


Figure 16. Glass chart and *N* scale.

One other point calls for examination before we can say that the preliminary investigation is complete. We must know whether pairs of glasses having the relation indicated by the numerical work are within the range of available types. Figure 16 shows a chart of the standard glasses made by Messrs. Chance Bros. with the ν's plotted logarithmically. A scale of values of ω has been added at the top of the diagram. Placed on it is a scale showing the values of N. The infinity mark on this scale is placed on the ν level corresponding to one of the glasses—in this case a flint of reciprocal index 0·60 (in the absence of a glass the corresponding point on the line of dense flints has been taken), and the value of N corresponding to any crown glass used in combination with this flint is then read off the scale. By comparing this chart with the values recorded in the table it is clear that suitable glasses are possible. The results of calculations on the preferred form of lens with different indices shows that a good choice of glasses is possible for constructing lenses of this type.

§ 6. SUBSEQUENT CALCULATIONS

So far as general exploration is concerned this investigation is now substantially completed. The next stage may either be the determination of the aberrational coefficients of the next order, still with undetermined glasses, or, as is more usual, the selection of probable glasses for a trial instrument, the insertion of suitable thicknesses in the lenses, and the introduction of small changes in the curvatures to take account of the effect of glass thickness. Experience with cemented lenses roughly comparable with those of the system portrayed in figure 14 leads one to expect that, at least so far as the central portion of the field is concerned, no particular difficulties will arise in making these alterations. This subsequent work is not considered here, for the purpose of this paper is to illustrate a method of optical prospecting developed at the National Physical Laboratory.

§ 7. ACKNOWLEDGMENT

The work described above was carried out in the Light Division of the National Physical Laboratory on behalf of the Ministry of Aircraft Production, by whose permission this paper is published.

APPENDIX

The formulae used in calculating the six aberrational coefficients are as follows :

$$a_1=\tfrac{1}{2}(1+N),\quad a_2=\tfrac{1}{2}(1-N),\quad a=a_1+a_2=1,$$

$$j_1=\frac{a_1}{1-\omega_1},\quad j_2=\frac{a_2}{1-\omega_2},\quad j=j_1+j_2,$$

$$k_1=j_1\omega_1=j_1-a_1,\quad k_2=j_2\omega_2=j_2-a_2,\quad k=k_1+k_2.$$

k, k_1 and k_2 are the total curvatures of the achromatic lens and of its crown and flint components respectively.

$$\varpi=a_1\omega_1+a_2\omega_2,\quad \alpha=-j_1j_2(\omega_1-\omega_2),\quad \beta=\alpha(\omega_1+\omega_2),\quad \tilde{\gamma}=2\beta-\alpha,$$

$$\gamma_1=2\beta j_1-\alpha k_1,\quad \gamma_2=-2\beta j_2+\alpha k_2,\quad \gamma+\gamma_1+\gamma_2=j^2+\alpha k_1-\alpha k_2.$$

When ρ is eliminated between the equations for coma and spherical aberration, the quantity $\alpha(1+\varpi)-\beta$ occurs. An independent check on part of the calculation is obtained by seeing that $-a_1a_2j(\omega_1-\omega_2)$ has the same value. The standard curvatures of the doublet are $\frac{1}{2}k$, $-\frac{1}{2}(k_1-k_2)$, $-\frac{1}{2}k$, and those of a divided triplet (say crown divided) are $\frac{1}{2}k$, $\frac{1}{2}(k_2-\theta k_1)$, $-\frac{1}{2}(k_2+\theta k_1)$, $-\frac{1}{2}k$. The time required for calculating these six quantities and the standard curvatures is between one and five minutes.

The formulae for the chromatic variations are

$$\Delta\varpi=\omega_1{}^2-\omega_2{}^2,\quad \Delta\alpha=k,\quad \Delta\beta=j(\omega_1{}^2+\omega_2{}^2),\quad \Delta\breve{\gamma}=2\Delta\beta-\Delta\alpha,$$

$$\Delta\gamma_1=2j^2\omega_2{}^2+k(k_1-2k_2),\quad \Delta\gamma_2=-2j^2\omega_1{}^2+k(2k_1-k_2),$$

$$\Delta(\gamma+\gamma_1+\gamma_2)=k(k_1-k_2).$$

They follow immediately from the principal formulae on putting

$$\Delta a_1=\Delta j_1=-\Delta a_2=-\Delta j_2=\epsilon$$

and taking the coefficients of ϵ.

DISCUSSION

Mr. G. A. Richmond. Mr. Smith's paper is certainly very interesting and instructive, and should be welcomed by lens designers as the basis for further investigation.

There are one or two remarks I should like to submit, but as I have not had any opportunity of doing any computing on Mr. Smith's preferred form, they must be treated as my own personal assumptions. From the construction of the lens system, it will be agreed that as far as central definition is concerned, a high degree of perfection appears to be obtainable, but it would be most interesting to have the full story of the off-axis aberrations for comparison with existing types of telephoto lenses. If the definition of Mr. Smith's lens at the edge of the field is no better than that of other types having less distortion, then I think his form will be somewhat at a disadvantage.

The condition of few glass-to-air surfaces is an advantage from the point of view of contrast, but I am not sure that it is advisable to cement components of such large diameter. After all, the adjacent surfaces can be well made, and the extra variables might enable the designer to obtain better general definition.

For reasons previously mentioned, I can only express the doubt regarding higher-order aberrations at the margin of the field, and if this doubt is justified—and I sincerely hope it is not—I shall feel that Mr. Smith has tied his hands unnecessarily by correcting each component in itself as well as cementing them.

In conclusion, I would point out that while it is recognized that distortion in long-focus reconnaissance lenses is of little consequence, cases may arise in which it must be considered.

Author's reply. The paper should be regarded as an illustration of one method of reaching a provisional design. In the earliest stages, when the number of degrees of freedom is potentially large, it seems reasonable to reduce these by assuming that very simple forms of construction are advantageous, and to find, under these conditions, what can be done to remove low-order aberrations. Having found any types of construction promising enough for fuller investigation, an increased number of variables will probably be required to control higher-order aberrations or to meet other conditions. There is nothing in the method described in this paper to prevent this. For example, the conditions involving ϕ as a free variable are not affected if the juxtaposed crown and flint surfaces differ in curvature, or if more than two components are used, or non-spherical surfaces are employed. Changes of this kind may be freely introduced and the new parameters given values which enable conditions involving both the first- and the second-order aberrations to be met. The region explored in the second stage will normally be limited in view of the conclusions reached in the first.

I agree with Mr. Richmond that in this particular case improvement is likely to be attained by increasing the number of glass-air surfaces. Some troubles will be sure to arise from refraction at the cemented surface of the positive component, and in practice, when a large relative aperture is required, at least four glass-air surfaces are likely to be used. While mentioning cemented surfaces, let me say that I realize the practical difficulties experienced in cementing large lenses successfully, and that these are increased by the extreme temperatures to which the lenses may be subjected. Nevertheless, cemented combinations have considerable advantages, and I hope ways of cementing lenses of large diameter will be thoroughly investigated in the future.

I am glad to have Mr. Richmond's view of the defects likely to be found in a lens of what I have called the preferred form. At the moment nothing more has been done than to ascertain that, while considerable higher-order aberrations are present, they are not of such a magnitude as to discourage the attempt to design a system based on this form. I hope it will be possible to make the comparison Mr. Richmond suggests between this and the customary forms of construction of telephoto lenses.

VARIATIONAL FORMULÆ IN OPTICS

By T. SMITH, F.R.S.,
Teddington

MS. received 8 *June* 1945

ABSTRACT. Formulae for calculating the effect in a complete optical instrument of small changes in the powers and separations of its component parts are collected together. The coefficients in these formulae need not be specially calculated, for they are already known if rays have been traced through the original system algebraically.

§ 1. INTRODUCTION

In recent papers Prof. M'Aulay and Mr. Cruickshank (1945) have given formulae for calculating the effect of small changes in optical systems, and have described their method of using them in working out the constructional specifications of optical instruments required by the Australian Armed Forces. They have based their work on Conrady's well-known book, and the fact that they have adopted his notation will no doubt be much appreciated by the many designers who use it. It is interesting that these authors, with no previous experience of optical designing, and presumably with no access to other works on technical optics, have realized the importance of branching out in this way. They have made useful additions to the normal computing routine, and it is rather surprising that formulae of this kind have been used so little, if at all, by many workers who rely on trigonometrical methods.

Expressions for the same purposes, constructed for use in conjunction with algebraic methods of lens calculation, have been in use for many years. To the writer a completely algebraic method seems to be altogether preferable to a trigonometric alternative. For the purposes now considered it yields the formulae desired more simply, and in application involves less calculation. Most of the formulae employed appeared originally in papers dealing with other problems, and it will probably be useful to collect them here. The notation used has been found convenient in many algebraic investigations. Conrady's and other trigonometric notations are not suitable.

Algebraic formulae for lens calculations can be constructed in very great variety. Those of a general character have common structural features which may not always be obvious. The formulae to be given depend on this structure and have a wide range of application. No difficulties are involved in the special applications, and the formulae are best developed in a general form. The essential property is that the characteristics of a symmetrical optical instrument can be expressed in terms of four functions which are regarded as the elements of a square matrix of order 2. When several lenses are arranged to form an instrument, the four functions of the combination are found from those of the component lenses by multiplying the corresponding matrices in their proper order. The determinant corresponding to each matrix has the value unity, so that only three of the functions are independent.

§2. MATRIX MULTIPLICATION

A brief explanation of matrix multiplication should perhaps be included for the benefit of some readers. A matrix is a collection of elements which are arranged in rows and columns and are identified by their positions. If $m_{p,q}$ is the element which lies at the intersection of the pth row and qth column of a matrix m, and $n_{q,r}$ is similarly the element which lies at the intersection of the qth row and rth column of a matrix n, the element of the product mn of the two matrices which lies in the pth row and rth column is by definition $\sum_q m_{p,q} n_{q,r}$, where the summation extends over all the columns of the factor m on the left and all the rows of the factor n on the right. (The number of rows in n is assumed to be equal to the number of columns in m.) It follows from this definition that in general the products mn and nm are different, so that matrices do not obey the commutative law for multiplication. With this exception they obey the elementary laws which hold in scalar algebra. For example, in forming the product of three matrices u, v, w, we may either form the product uv first and post-multiply it by w, or first form the product vw and pre-multiply it by u.

The matrices to be considered here have two rows and two columns, and the standard notation employed is

$$m = \begin{pmatrix} B & D \\ A & C \end{pmatrix}.$$

From the law of multiplication, if $m_1 m_2 = m_{1,2}$,

$$A_{1,2} = A_1 B_2 + C_1 A_2, \qquad C_{1,2} = A_1 D_2 + C_1 C_2,$$
$$B_{1,2} = B_1 B_2 + D_1 A_2, \qquad D_{1,2} = B_1 D_2 + D_1 C_2.$$

This law of multiplication is identical with one of the common formulae for multiplying determinants, and as the determinant corresponding to any elementary matrix has the value unity, it follows that in all cases

$$BC - AD = 1.$$

If E denotes the matrix

$$\begin{pmatrix} 1 & \cdot \\ \cdot & 1 \end{pmatrix}$$

(where the dots represent zero elements), the law of multiplication gives

$$Em = m = mE,$$

so that E is the analogue of the scalar number 1 and commutes with all matrices m.

If m is pre- and post-multiplied by

$$\begin{pmatrix} C & -D \\ -A & B \end{pmatrix},$$

the product in both cases is E. This matrix therefore commutes with m and has the properties of an inverse with respect to m; it is appropriately denoted by m^{-1}. It is easy to show that if $m_1 m_2 = m_{1.2}$, then $m_{1.2}^{-1} = m_2^{-1}\, m_1^{-1}$, i. e. the order of the factors in the inverse of a product is the reverse of that in the original product.

The only further point that appears to need explanation is the meaning of a matrix equation such as $m = n$. This equation implies that any element of m is equal to the correspondingly placed element of n; it is thus equivalent to several scalar equations.

§ 3. THE MATRIX ELEMENTS

The events to be recorded in order to represent the passage of light through an optical instrument are the changes of direction brought about by reflexion or refraction at a surface, and the displacements on proceeding from one surface to the next. The instrumental matrix m is built up from matrix factors which represent alternately events of these two types. For the directional changes, the factor matrix has D zero and for displacements A is zero. In most cases both B and C are equal to unity in the elementary factors, and in all the product BC is equal to unity. Different systems can be constructed with somewhat different meanings for the four functions, but it may be assumed here that A represents the power of a surface or of an instrument, and $-D$ the distance along a ray between successive surfaces divided by the refractive index, or, for a complete instrument, D is the apparent distance between the extreme points for which the calculation is made. B and C are angular or linear magnifications. Assuming that the matrix relates to a (possibly) skew ray, the precise meanings can be seen in all cases from the equation

$$\begin{pmatrix} \xi' & -x' \\ \eta' & -y' \end{pmatrix} = \begin{pmatrix} \xi & -x \\ \eta & -y \end{pmatrix} \begin{pmatrix} B & D \\ A & C \end{pmatrix},$$

where x, y are the co-ordinates of the initial, and x', y' of the final point of the ray, measured in planes normal to the axis of symmetry, and ξ, η and ξ', η' are the products of the corresponding direction cosines and the refractive indices of the first and last media. The same A, B, C, D clearly determine the sagittal focal lines. In all the elementary factors B and C are both equal to unity.

If the matrix relates to tangential focal lines, the displacement factor is the same as before, but the elementary refraction factor is

$$\begin{bmatrix} \dfrac{\cos\phi}{\cos\phi'} & \cdot \\[2ex] \dfrac{A}{\cos\phi\,\cos\phi'} & \dfrac{\cos\phi'}{\cos\phi} \end{bmatrix}$$

in terms of the sagittal A, where ϕ and ϕ' are the angles of incidence and refraction.

The significance of the functions in m in this case is shown by the equation

$$(\mu'\delta\psi' \quad -\delta \mathfrak{h}') = (\mu\delta\psi \quad -\delta \mathfrak{h})m,$$

where $\delta \mathfrak{h}$ is the length of the perpendicular from the initial point of the principal ray of the pencil to another ray inclined to it at the angle $\delta\psi$, and accented quantities relate to the final values. For example, if u and v are the distances of an object point and of its focal line, the relation they satisfy is

$$\left(1 \quad -\frac{u}{\mu}\right) m \begin{bmatrix} -\dfrac{v}{\mu'} \\ 1 \end{bmatrix} = 0 \; ;$$

this is merely another way of saying that, when the end points are the object and image positions, $\delta \mathfrak{h}'$ depends on $\delta \mathfrak{h}$ but not on $\delta\psi$, i.e. $D=0$.

§ 4. VARIATIONAL FORMULAE

The formulae to be considered express the changes in the system elements A, B, C, D due to small changes in the construction of the system, such as changes of refractive index, curvature, or axial separation. As the changes are small, their effects are directly additive, and it is only necessary to consider a single alteration to obtain the expressions required. In general, the part to be altered will be preceded by an unaltered portion of the system, and followed by another unaltered part. We therefore write

$$m = m_1 \, . \, m_2 \, . \, m_3,$$

where the part to be altered carries the suffix 2. On making the changes we clearly have the variational equation

$$\delta m = m_1 \, . \, \delta m_2 \, . \, m_3. \qquad \ldots\ldots(1)$$

It is sometimes preferred to use equations derived from this by pre- or post-multiplication by m^{-1}, i. e. to use

$$m^{-1} \, . \, \delta m = m_{2,3}^{-1} \, . \, \delta m_2 \, . \, m_3. \qquad \ldots\ldots(2)$$

or

$$\delta m \, . \, m^{-1} = m_1 \, . \, \delta m_2 \, . \, m_{1,2}^{-1}. \qquad \ldots\ldots(3)$$

The most important cases are when the system 2 consists only of a single surface or a single separation, and the change is either an alteration in A_2 or in D_2. For a change of power, equation (1) may be written

$$\begin{pmatrix} \delta B & \delta D \\ \delta A & \delta C \end{pmatrix} = \delta A_2 \begin{pmatrix} D_1 \\ C_1 \end{pmatrix} (B_3 \quad D_3), \qquad \ldots\ldots(4)$$

and for an altered separation

$$\begin{pmatrix} \delta B & \delta D \\ \delta A & \delta C \end{pmatrix} = \delta D_2 \begin{pmatrix} B_1 \\ A_1 \end{pmatrix} (A_3 \quad C_3). \qquad \ldots\ldots(5)$$

Thus the change δA in the power of the system due to the change of power δA_2 of the surface 2 is $C_1 B_3 \delta A_2$, and that due to the separation change δD_2 is $A_1 A_3 \delta D_2$.

It will be noted that relations derived from equation (1) involve a knowledge of the functions relating to both the preceding and following parts of the instrument; in other words, the functions require to be calculated from each end of the system. Duplicating the numerical results in this way is an excellent check on the accuracy of the calculations, for the intermediate quantities are all different. But in many cases it will be desired to avoid this, and equations included in (2) or (3) are then appropriate. For (2) the calculated values are found from a reverse trace, and for (3) from a direct trace. In the most general case the relations are

$$A^2\delta\left(\frac{B}{A}\right) = -B_{2,3}B_3\delta A_2 + A_{2,3}B_3\delta B_2 - B_{2,3}A_3\delta C_2 + A_{2,3}A_3\delta D_2,$$

$$D^2\delta\left(\frac{C}{D}\right) = D_{2,3}D_3\delta A_2 - C_{2,3}D_3\delta B_2 + D_{2,3}C_3\delta C_2 - C_{2,3}C_3\delta D_2,$$

$$A^2\delta\left(\frac{C}{A}\right) = -C_1C_{1,2}\delta A_2 - A_1C_{1,2}\delta B_2 + C_1A_{1,2}\delta C_2 + A_1A_{1,2}\delta D_2,$$

and

$$D^2\delta\left(\frac{B}{D}\right) = D_1D_{1,2}\delta A_2 + B_1D_{1,2}\delta B_2 - D_1B_{1,2}\delta C_2 - B_1B_{1,2}\delta D_2.$$

Most frequently we require to know the changes due to the alteration of a single element. If only a separation is altered we insert zero values for δA_2, δB_2, δC_2 and note that $A_{1,2} = A_1$, $B_{1,2} = B_1$, $A_{2,3} = A_3$, $C_{2,3} = C_3$. For forward traces the formulae are then

$$A^2\delta\left(\frac{C}{A}\right) = -\delta\left(\frac{C}{D}\right) + 2C\delta\left(\frac{1}{D}\right) - C^2\delta\left(\frac{B}{D}\right) = A_1^2\delta D_2,$$

$$A\left\{\delta\left(\frac{1}{A}\right) - B\delta\left(\frac{C}{A}\right)\right\} = -D\left\{\delta\left(\frac{1}{D}\right) - C\delta\left(\frac{B}{D}\right)\right\} = -A_1B_1\delta D_2,$$

$$\delta\left(\frac{B}{A}\right) - 2B\delta\left(\frac{1}{A}\right) + B^2\delta\left(\frac{C}{A}\right) = -D^2\delta\left(\frac{B}{D}\right) = B_1^2\delta D_2,$$

and for backward traces are

$$A^2\delta\left(\frac{B}{A}\right) = -\delta\left(\frac{B}{D}\right) + 2B\delta\left(\frac{1}{D}\right) - B^2\delta\left(\frac{C}{D}\right) = A_3^2\delta D_2,$$

$$A\left\{\delta\left(\frac{1}{A}\right) - C\delta\left(\frac{B}{A}\right)\right\} = -D\left\{\delta\left(\frac{1}{D}\right) - B\delta\left(\frac{C}{D}\right)\right\} = -A_3C_3\delta D_2,$$

$$\delta\left(\frac{C}{A}\right) - 2C\delta\left(\frac{1}{A}\right) + C^2\delta\left(\frac{B}{A}\right) = -D^2\delta\left(\frac{C}{D}\right) = C_3^2\delta D_2.$$

Similarly if

$$m_2 = \begin{pmatrix} 1 & \cdot \\ A_2 & 1 \end{pmatrix},$$

so that only A_2 is changed, $C_{1,2} = C_1$, $D_{1,2} = D_1$, $B_{2,3} = B_3$, $D_{2,3} = D_3$ and the equations are

$$A^2\delta\left(\frac{C}{A}\right) = -C_1^2\delta A_2, \quad A\left\{\delta\left(\frac{1}{A}\right) - B\delta\left(\frac{C}{A}\right)\right\} = C_1D_1\delta A_2,$$

$$\delta\left(\frac{B}{A}\right) - 2B\delta\left(\frac{1}{A}\right) + B^2\delta\left(\frac{C}{A}\right) = D_1^2\delta A_2,$$

$$A^2\delta\left(\frac{B}{A}\right) = -B_3^2\delta A_2, \quad A\left\{\delta\left(\frac{1}{A}\right) - C\delta\left(\frac{B}{A}\right)\right\} = B_3D_3\delta A_2,$$

$$\delta\left(\frac{C}{A}\right) - 2C\delta\left(\frac{1}{A}\right) + C^2\delta\left(\frac{B}{A}\right) = -D_3^2\delta A_2,$$

where the expressions in terms of A may be replaced by their equivalents in terms of D already given. It will be noted that use has been made of equations (4) and (5) as well as of (2) and (3) to construct sets of formulae for one-way operation.

If

$$m_2 = \begin{bmatrix} \dfrac{\cos\phi}{\cos\phi'} & \cdot \\ A_2 & \dfrac{\cos\phi'}{\cos\phi} \end{bmatrix},$$

and only A_2 is changed, the right sides of the forward formulae just given each contain the additional factor $\dfrac{\cos\phi'}{\cos\phi}$ and the backward formulae the factor $\dfrac{\cos\phi}{\cos\phi'}$.

Many applications of these relations concern the determination of changes of power or separation required to bring about desired alterations in the functions for the complete instrument rather than the converse problem so far considered. The changes are determined by the solution of a number of linear simultaneous equations, and it is useful to have some guide to aid in selecting the most suitable elements for alteration. To ensure that all changes are small, the determinant of the coefficients in the equations of the elements to be modified must not be small. As a rule functions relating to a number of rays will be involved, and the type of construction enters into the problem. But some guidance can be gained from the coefficients for a single ray. A typical case is when the focal length and back focal length, i. e. $1/A$ and C/A, are to be adjusted. We can then regard such quantities as $C_1\delta A_2$ or $A_1\delta D_2$ as independent variables, with C_1 and D_1 or A_1 and B_1 as coefficients. Changing for convenience to a double suffix solution, the matrix relations give for an instrument with n spaces

$$\begin{vmatrix} C_{1,p-1} & D_{1,p-1} \\ C_{1,q} & D_{1,q} \end{vmatrix} = D_{p,q} = \begin{vmatrix} B_{q+1,n} & D_{q+1,n} \\ B_{p,n} & D_{p,n} \end{vmatrix},$$

and for one with n surfaces

$$\begin{vmatrix} A_{1,p-1} & B_{1,p-1} \\ A_{1,q} & B_{1,q} \end{vmatrix} = -A_{p,q} = \begin{vmatrix} A_{q+1,n} & C_{q+1,n} \\ A_{p,n} & C_{p,n} \end{vmatrix},$$

so that in one case the apparent distance between the surfaces to be altered should be large, and in the other the power of the system between the separations to be changed should be large. If these quantities are small it will usually only be practicable to satisfy one of the selected conditions instead of both.

If three conditions are imposed instead of two, none of the three apparent distances, or none of the intervening powers, may be small. This is evident if, for instance, we note that

$$\begin{vmatrix} C^2_{1,p-1} & C_{1,p-1}D_{1,p\ 1} & D^2_{1,p-1} \\ C^2_{1,q} & C_{1,q}D_{1,q} & D^2_{1,q} \\ C^2_{1,r} & C_{1,r}D_{1,r} & D_1{}^2{}_{,r} \end{vmatrix} = D_{p,q}D_{q+1,r}D_{p,r}.$$

In his paper already mentioned, Prof. M'Aulay (1945) puts forward approximate formulae for focal displacements. It may be well to point out that I gave the exact law which applies in these cases many years ago (1922/23), and also the form of the sine condition for oblique rays (1929/30), which is given by the two equations

$$Axx' - x\xi + x'\xi' + D\xi\xi' = 0,$$
$$Ayy' - y\eta + y'\eta' + D\eta\eta' = 0.$$

Since, following Conrady, this paper ascribes the relation $x\xi = x'\xi'$ to Lagrange, it should be pointed out now, when it is being repeated in these *Proceedings*, that this is one of several historical inaccuracies to be found in Conrady's book.

REFERENCES

CRUICKSHANK, F. D., 1945. "A system of transfer coefficients for use in the design of lens systems." *Proc. Phys. Soc.* **57,** 350, 362, 419, 426, 430.

M'AULAY, A. L., 1945. "A transfer method for deriving the effect on the image. . . . from ray changes produced at a given surface." *Proc. Phys. Soc.* **57,** 435.

M'AULAY, A. L. and CRUICKSHANK, F. D., 1945. "A differential method of adjusting the aberration of a lens system". *Proc. Phys. Soc.* **57,** 302.

SMITH, T., 1922/3. "The optical cosine law". *Trans. Opt. Soc.* **24,** 31.

SMITH, T., 1929/30. "The general form of the Smith-Helmholtz equation." *Trans. Opt. Soc.* **31,** 241.

THE ELECTRON MICROSCOPE AND ITS APPLICATIONS

BY F. W. CUCKOW,
Teddington

Lecture-Survey delivered 23 *February* 1945

THE speaker mentioned the various forms of electron microscope—those producing images of self-emitting objects; those producing point sources of electrons which were then used to project shadows of nearby objects; scanning microscopes in which the object was scanned by a fine electron probe and images of the object obtained in terms of the secondary emission; and finally, a type very similar in its component parts to an ordinary light microscope, the transmission electron microscope. The greater part of the lecture was devoted to this last type.

After tracing events leading to the development of electron microscopes, the resolving powers to be expected were discussed together with the limitations set by existing electron lenses.

Slides were shown in illustration of the various practical electron microscopes already constructed in America, Belgium, Britain, Canada and Germany, including the commercial instruments produced in the first and last named countries.

It was pointed out that the maximum "aperture" of existing electron microscopes was only about one-thousandth of that attained in high-power light microscopes on account of the severe spherical aberration affecting electron lenses.

Associated with this small "aperture", however, was a far greater depth of focus.

The techniques for specimen preparation and the applications of the instrument were discussed together. New information has been obtained on the fine structure of diatoms, which are found to be large enough to be mounted over the holes in a fine wire mesh.

Most biological material and specimens such as fine powders were mounted on films of nitrocellulose or other plastic mounted in turn on the fine wire mesh. The plastic films have to be of the order of one-millionth of an inch in thickness to avoid undue scattering of the electron beam. Filter-passing viruses, and powders, such as paint pigments, ceramic materials and fillers, and smoke particles were mounted in this manner.

The transmission type of electron microscope could not deal directly with massive specimens requiring surface examination, but such specimens could sometimes be examined indirectly by "replica" techniques ; these had been used in studying the microstructure of teeth and metals.

Finally, the migratory bands, produced in electron micrographs by electron diffraction from the atomic planes of crystalline speciments, were mentioned, and also electron diffraction "past a straight edge" giving a banded pattern exactly comparable with that obtained in the well-known experiment with light.

DISCUSSION

Mr. C. G. **Wynne**. It is well known that the resolving power of a lens may be considerably increased, at the expense of some loss of contrast, by the use of an annular aperture. This effect would appear to be particularly suitable for use in the electron microscope lenses, with their heavy spherical aberration, and I should be interested to know whether it has been tried. In the presence of uncorrected spherical aberration the obscuring of a central part of the beam without any increase in the maximum angular aperture of the beam (2α in Mr. Cuckow's lecture) would give increased resolution and a reduction in the effective aberration, together with reduced intensity, and this would not be conditional on accurate centring of the annular aperture, which it is suggested might be difficult. If this possible difficulty of centring could be overcome (for example by using the annular aperture at such a position that it could be imaged in the same way as the source itself is imaged for centring) then the improved aberration would permit a small increase of α, with a further gain in resolving power. As α is still further increased, the effective spherical aberration may be controlled by reducing the width of the annulus; since the width of a constant-area annulus decreases proportionally to the first power of the reciprocal of its increasing mean radius, while the aberration will increase as higher powers, large increases of α will necessitate a rapid decrease of aperture area and hence of intensity. It would appear that this will be the limiting factor in the process, so that it would be interesting to know to what extent the intensity of the source could be increased to allow the fullest scope for the method.

Mr. R. N. **Crooks**. Members will probably be interested to hear that a number of electron microscopes are now being made in this country, and the firm concerned is the one that made the first instrument illustrated by Mr. Cuckow. Although production was suspended until recently owing to more pressing war needs, there is no reason to anticipate that the new instruments will not give as good a performance as the ones now being manufactured in the United States. The first instrument is expected to be ready about the middle of this year.

Mr. M. E. **Haine**. I would like to say a few words concerning the history of the development of the electron microscope.

Early work was carried out mainly in Germany by Knoll and Ruska, these workers having constructed their first transmission-type instrument about 1931. Further work was done in the early thirties in Germany by the same workers and others, and by Marton in Brussels.

In 1935, in this country, work was commenced on the construction of an instrument to the design of Professor Martin, of the Imperial College. The instrument was constructed in the Research Department of Metropolitan-Vickers, and on completion was transferred to the Imperial College, and several years later to the N.P.L. The instrument never got past the stage of being an experimental model, although it ultimately gave a good performance. Further work in this country was held up by the advent of the war until quite recently: but in the late thirties Canadian workers at Toronto, Prebus and Hillier, developed a first-class instrument.

By 1938 commercial instruments were available in Germany, and shortly after, R.C.A. of America, having taken Marton from Belgium and Hillier from Toronto on their staff, produced their model B, which is the instrument mainly described in this paper.

In the last year, work has been recommenced in this country, and a war-time "Utility" electron microscope is now on the market. This instrument was designed and is being produced by the Research Department of Metropolitan-Vickers in Manchester.

With reference to a previous speaker's remarks concerning the use of an annular aperture in the objective stage of the microscope, it should be noted that the width of such an aperture must be altered by an amount inversely proportional to the square of the radius of the annulus in order to keep the spherical aberration below the diffraction error. The area of the aperture will then be inversely proportional to the radius. This will result in a reduction in intensity, unless the aperture is placed in the condenser lens, in which case an increased current density might be utilized without over-heating the specimen. However, a further limitation now creeps in, due to the scattering of the beam in the specimen, which now produces r^3 times the aberration produced in the circular-aperture case.

Mr. H. H. HOPKINS. It has been suggested by Mr. C. G. Wynne that the use of a narrow annular aperture stop for the electron beam would serve to reduce the effects of spherical aberration, while having advantages over the small circular stop at present used in giving increased resolution.

In the case of light waves, Rayleigh showed that the use of an annular stop gave a narrower central maximum than the usual Airy-disc distribution, together with increased intensity in the diffraction rings. He advocated the use of such a stop in solar observations, whereby the illumination could be reduced without impairing the resolving power —in fact, this is enhanced. The problem has been treated in detail by G. C. Steward.

Mr. T. Smith has said that the effects of the annular aperture in the case of an off-axis object point have not been investigated. I would like to correct this impression. G. C. Steward has given a full treatment of the case, both for a spherical wave and for a wave suffering from first-order coma. Curves are drawn, and it is concluded "... this illustrates a decided gain in resolving power over the full aperture case". It would, in any case, be surprising to observe a sharp discontinuity in these effects when the object-point is moved off-axis by a distance of first-order magnitude.

The same effects are well known to microscopists in the use of so-called oblique illumination, which is, of course, a common practice.

If the electron beam had a circular stop 0·10 inch in diameter instead of the diameter of 0·02 inch which has been mentioned, the resolving power would, theoretically, be increased five times. An annular stop—if, in this case, an electron beam behaves the same as light waves—would give an even greater increase. It would, I am sure, be most interesting to know whether this is the case in practice.

REFERENCES

RAYLEIGH, Lord, 1887/92. *Sci. Pap.* iii. p. 90.

STEWARD, G. C., 1925. "The Symmetrical Optical System", *Cambridge Tracts*, no. 25, p. 88 *et seq.*, and *Philos. Trans.* A, **225**.

AUTHOR'S reply. I find that most of the points raised have been dealt with by other speakers. No references have been seen to attempts to use annular apertures, and it seems to be a matter which might well repay practical attention. One possible difficulty would be the apparent necessity for accurately centring the aperture so that all rays passed through the system at a common inclination to the axis. I cannot agree with Mr. Wynne that the " increased resolution . . . would not be conditional on accurate centring " even for such small-angle beams as are now used, for if the centre is blocked out of a non-axial beam produced by a de-centred aperture, the residual rays pass through the system at a variety of inclinations to the axis, whereas for an axial beam the residual rays pass through at a common inclination. The partly blocked out non-axial beam might, therefore, be expected to give a less good image on account of spherical aberration than the partly blocked out axial beam.

The advent of a commerically produced British instrument is awaited with much interest. Messrs. Metropolitan-Vickers Electrical Company are to be congratulated on proceeding with this work under the conditions which have prevailed in this country of late.

FLAT-FIELDED SINGLET APLANATS

By C. R. BURCH, F.R.S.,
H. H. Wills Physical Laboratory, University of Bristol

MS. received 15 *July* 1945

ABSTRACT. The design of flat-fielded aspheric aplanatic singlets is discussed by means of the " plate diagram ", and two flat-fielded singlets of $n=1{\cdot}525$ are designed. Flat-fielded mirror pairs are discussed, and it is shown that a sufficient condition for Seidel field-flatness is that object and image should lie on planes tangent to the mirror surfaces. Some formulae relating to reflecting microscope design are given.

§ 1. INTRODUCTION

SINGLET lenses can be designed to give flat-fielded anastigmatic images of distant objects, in the extended paraxial region, provided the refractive index exceeds a certain critical value $n=1{\cdot}602\ldots.$ Both surfaces must be aspherized unless $n=1{\cdot}618\ldots.$, when the convex face may be spherical (Burch, 1943). L. C. Martin (1944), applying Chrétien's aplanat theory to the " critical case ", $n\doteq1{\cdot}602\ldots.$, has obtained design formulae valid outside the extended paraxial region, and has shown by ray-tracing that the resulting lens, at f/5, is well corrected (as a monochromat) over a semi-field of 7° 30′. The lateral colour error of these lenses may be anulled by a diaphragm placed a short distance in front of the front surface. The longitudinal colour error of course remains. To reduce this error we may prefer to select a value of n less than $1{\cdot}602\ldots.$, though this means sacrificing exact flat-fielded anastigmatism. So we are led to discuss the theory of flat-fielded aspheric aplanatic singlets. This may conveniently be done by the method of the plate diagram described in the previous paper (Burch, 1943).

§ 2. GENERAL THEORY OF APLANATIC SINGLETS

We shall use the notation of the previous paper : let the singlet have a front surface of paraxial radius ρ (positive if convex), back surface of paraxial radius R (positive concave), thickness d and refractive index n. It is necessary to define the position of the object in some manner, and we do so by means of the position of the intermediate image, because this is referable with equal analytical simplicity to either surface. Let this image be formed in glass at a distance L inside the pole of the front surface (see figure 1). We can regard this singlet as an anastigmat to which there have been added four aspheric plates, two of these being the figurings on the ρ- and R- surfaces respectively, and the other two the negatives of the "Schmidt" plates which would be necessary at the ρ and R centres to anastigmatize the refractions at the ρ and R surfaces if these were spherical. For the more convenient summation of the effects of these aspheric plates we create an artificial *star-space* (i.e. a space in which the object is at infinity, "among the stars") by allowing a hypothetical anastigmat of unit focal length (in refractive index n) to act on the intermediate image. In this star-space we draw the *plate diagram*. Let m_1, m_2 be the strengths in star-space of the missing anastigmatizing plates of the ρ- and R-surfaces and m_3, m_4 the strengths of the ρ- and R-figurings. Let the distances in star-space between these plates be as shown in figure 1.

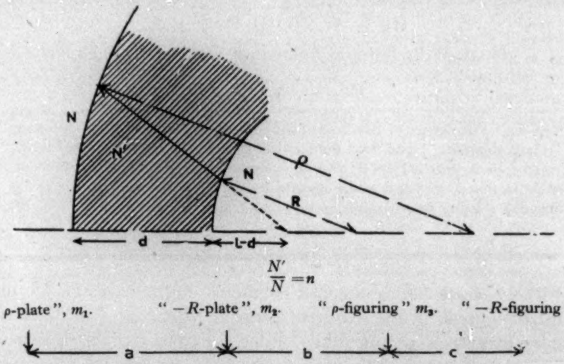


Figure 1. Two-surface system and its plate diagram.

It will suffice to quote from the previous paper (Burch, 1943) the formulae for these strengths and distances.

We have then

$$\left.\begin{aligned} a&=\frac{(R+d-\rho)}{(\rho-L)(R+d-L)} & (a+b)&=\frac{\rho}{L(\rho-L)} \\ b&=\frac{(R+d)}{L(R+d-L)} & (b+c)&=\frac{R}{(L-d)(R+d-L)} \\ c&=\frac{d}{L(L-d)} & (a+b+c)&=\frac{(\rho-d)}{(L-d)(\rho-L)} \end{aligned}\right\} \qquad \ldots\ldots(1)$$

ρ-plate strength,

$$m_1=\frac{H^4(n-1)n^2L^2(L-\rho)^2}{8\rho^3}\left\{\frac{L-\left(\frac{n+1}{n}\right)\rho}{L}\right\}, \text{ retardation lacking;}$$

R-plate strength

$$m_2 = \frac{-H^4(n-1)n^2(L-d)^2(L-d-R)^2}{8R^3}\left\{\frac{(L-d)-\left(\frac{n+1}{n}\right)R}{L-d}\right\},\ \text{retardation lacking,} \qquad \ldots\ldots(2)$$

H being the zonal height in the artificial star-space.

Now, if in any plate diagram we equilibrate the zeroth and first moments about any pivot-point of a plate of strength m, retardation lacking, by means of two plates, one of retardation $\frac{m\alpha}{\alpha+\beta}$ distant β to the left of the $-m$ plate and one of retardation $\frac{m\beta}{\alpha+\beta}$ distant α to the right of it, we shall provide a surplus of second moment (about any pivot point) of amount $\alpha\beta$. Accordingly in aplanatizing the singlet, i.e. in annulling by suitable choice of the figurings m_3 and m_4 the zeroth and first moments of the plates m_1 and m_2, we shall provide an excess of second moment consisting of two contributions, associable with the ρ- and R-plates respectively.

For the ρ-plate

$$\left.\begin{array}{l} \alpha_\rho=(a+b+c);\ \beta_\rho=-(a+b);\ [m\alpha\beta]_\rho=-(a+b)(a+b+c)m_1. \\ \text{For the } R\text{-plate} \\ \alpha_R=(b+c)\ ;\qquad \beta_R=-b\ ;\qquad [m\alpha\beta]_R=-b(b+c)m_2. \end{array}\right\}\ldots\ldots(3)$$

Thus the second moment of retardation of the aplanatic singlet is

$$\frac{-H^4(n-1)n^2}{8L(L-d)}\left[(\rho-d)\frac{L}{\rho}\left\{\frac{L}{\rho}-\frac{n+1}{n}\right\}-(R+d)\frac{(L-d)}{R}\left\{\frac{L-d}{R}-\frac{n+1}{n}\right\}\right]. \qquad \ldots\ldots(4)$$

It is easy to show that plate-diagram second moments are independent of the focal length of the diagram, and are proportional to the square of the refractive index of the star-space. Eight times the coefficient of second moment of retardation of a plate diagram drawn in air, when multiplied by the square of image height, gives the negative astigmatic interfocal distance of the image. Accordingly at image height h the aplanatic single has

Negative A.I.D. =

$$\frac{-h^2(n-1)}{L(L-d)}\left[(\rho-d)\frac{L}{\rho}\left\{\frac{L}{\rho}-\frac{n+1}{n}\right\}-(R+d)\frac{(L-d)}{R}\left\{\frac{L-d}{R}-\frac{n+1}{n}\right\}\right]. \qquad \ldots\ldots(5)$$

It also has Petzval depth, given by

$$\frac{h^2(n-1)}{2n}\left(\frac{1}{\rho}-\frac{1}{R}\right). \qquad \ldots\ldots(6)$$

Let us equate (5) to κ times (6). This gives the condition that the astigmatic interfocal distance should be κ times that needed to make the surface of best definition flat, when the singlet is rendered aplanatic by figuring

$$\frac{L}{\rho}\left[\left\{\frac{L}{\rho}-\frac{n+1}{n}\right\}(\rho-d)+\frac{\kappa(L-d)}{2n}\right]=\frac{L-d}{R}\left[\left\{\frac{L-d}{R}-\frac{n+1}{n}\right\}(R+d)+\frac{\kappa L}{2n}\right], \qquad \ldots\ldots(7)$$

which differs from the anastigmatism condition obtained in the previous paper only by the inclusion of the terms $\frac{\kappa(L-d)}{2n}$, $\frac{\kappa L}{2n}$.

If we fix the front radius, ρ, and the object distance, L, and satisfy (7), one of the two variables, R and d, remains at our disposal, and we may use it to annul either of the two figurings. We can achieve aplanatism with only one figured surface if, and only if, the combined first moments of the R- and ρ-plates about that surface are zero.

Thus the condition for aplanatism when the front surface only is figured is $m_1/m_2 = -a/(a+b)$, i.e.

$$\frac{L-\rho}{\rho}\left[\frac{L}{\rho}-\frac{n+1}{n}\right] = -\frac{(R+d)}{R}\frac{(R+d-L)}{R}\frac{(L-d)}{L}\left[\frac{L-d}{R}-\frac{n+1}{n}\right], \quad \ldots\ldots(8)$$

and the expression for the negative astigmatic interfocal distance then reduces to

$$\frac{-h^2(n-1)}{\rho}\left[\frac{L}{\rho}-\frac{n+1}{n}\right]\left[\frac{R+d-\rho}{R+d-L}\right]. \quad \ldots\ldots(9)$$

When the back surface only is figured, the condition for aplanatism is

$$\frac{(R+d-L)}{R}\left[\frac{L-d}{R}-\frac{n+1}{n}\right] = -\frac{(\rho-d)}{\rho}\frac{(L-\rho)}{\rho}\frac{L}{(L-d)}\left[\frac{L}{\rho}-\frac{n+1}{n}\right] \quad \ldots\ldots(10)$$

and expression (5) reduces to

$$\frac{-h^2(n-1)}{R}\left[\frac{L-d}{R}-\frac{n+1}{n}\right]\left[\frac{R+d-\rho}{\rho-L}\right]. \quad \ldots\ldots(11)$$

§ 3. SPECIAL CASE: OBJECT AT INFINITY

We shall now confine our attention to the case when the object is at infinity, that is, we set $L = n\rho/(n-1)$. Let us set

$$x = d/\rho\,; \quad y = \frac{L-d}{R}\,; \quad z = R/\rho = \frac{n-(n-1)x}{(n-1)y}, \quad \ldots\ldots(12)$$

and the condition that the astigmatism of the aplanat should be κ times that needed to flatten the field of best definition—equation (7)—reduces to

$$\begin{aligned} &z^2\left[x\left\{(n+1)\frac{(n-1)^2}{n}+\kappa\frac{(n-1)}{2}+1\right\}-n^2-\frac{n\kappa}{2}\right] \\ &+z\left[x^2(n-1)^2\frac{(2n+1)}{n}-x(n-1)\left(3n+1+\frac{\kappa}{2}\right)+n^2-\frac{n\kappa}{2}\right] \\ &+x[n-(n-1)x]^2=0. \end{aligned} \quad \ldots\ldots(13)$$

When κ is fixed, equation (13) defines a curve on the (x, z) diagram passing through all the points (x, z), representing singlet aplanats having field-flattening coefficient κ for distant objects. Two cases, $\kappa = 1$ (the flat-fielded aplanats; curve E) and $\kappa = 0$ (the anastigmats; curve F), are illustrated in figure 2, which has been computed for $n = 1{\cdot}525$. We may conveniently show on the same diagram the two curves, A, of the aplanats which only require figuring on the front surface, and B, of those which only require figuring on the back surface.

Curve A is obtained from equation (8), which reduces to

$$(n-1)^2(1-y)(1/n+1-y)[n-(n-1)x(1-y)]=1, \qquad \ldots\ldots(14)$$

while curve B is obtained from equation (10), which reduces to

$$y=1+\frac{1}{2n}\pm\sqrt{\frac{1}{4n^2}+\frac{1-x}{(n-1)^2+n-(n-1)x}}, \qquad \ldots\ldots(15)$$

y being defined by (12).

The lenses of curve A have negative A.I.D. given by

$$\frac{-h^2}{n\rho}\left(\frac{z+x-1}{z+x-\dfrac{n}{n-1}}\right) \qquad \ldots\ldots(16)$$

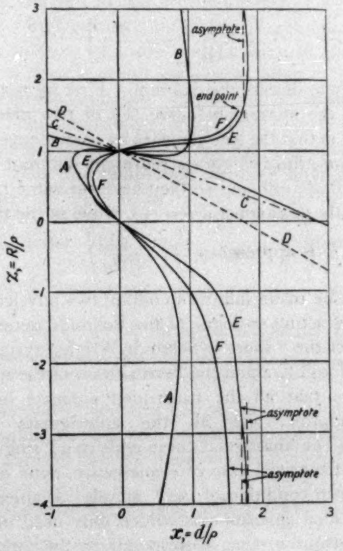


Figure 2.

from (9), while from (11) lenses of curve B have negative A.I.D. of

$$+\frac{h^2(n-1)}{z^2\rho}\left[n-(n-1)x-\left(\frac{n^2-1}{n}\right)z\right][z+x-1] \qquad \ldots\ldots(17)$$

For the sake of completeness we notice in the general case, when aplanatism is achieved by figuring both faces, the lens, whether flat-fielded or not, has a negative A.I.D. of amount

$$\frac{-h^2(n-1)}{n[n-(n-1)x]}\left[1-x-(z+x)\left(\frac{n-(n-1)x}{z}\right)\left(\frac{n-(n-1)x}{z}-\frac{n^2-1}{n}\right)\right].$$

$$\ldots\ldots(18)$$

We may conveniently add to the diagram the *afocal line*, C, the equation of which is

$$z+\left(\frac{n-1}{n}\right)x=1. \qquad \ldots\ldots(19)$$

The focal length of any singlet is obtained by dividing $R(=z\rho)$ by $(n-1)$ times the height above the afocal line of the representative point. Finally, we add the line D, which represents those singlets which are laterally achromatic when used as eyepieces with an objective of very great tube-length : the equation of this line is

$$z+\left(\frac{n^2-1}{n^2}\right)x=1, \qquad \ldots\ldots(20)$$

and the clear distance of the eyepoint of this eyepiece in front of the ρ-surface is

$$\frac{d}{(n-1)x^2}\left[x-\frac{n^2}{n-1}\right]. \qquad \ldots\ldots(21)$$

We are now ready to discuss the diagram. First we notice that the flat-field curve (E) intersects the anastigmat curve (F) in two places. One of these, $x=0$, $z=1$, is trivial, in that the lens vanishes by coincidence of the surfaces. At the second intersection, the two curves represent different lenses. Both share the properties $R/\rho=0$, $d/\rho=0$, and so they have the same representative point : they differ in that in the anastigmat $d/R=-1$, while in the flat-fielded lens, as we approach this point, d/R approaches $-\left(1+\frac{1}{2n}\right)$.

ρ of course must be made infinite to obtain two physically realizable lenses. The former is the refracting analogue of the Schmidt camera : the latter is the refracting analogue of the "short" (Schmidt-Wright) system, in which enough astigmatism is introduced to annul the Petzval term of the mirror.

Next we observe that all the flat-fielded aplanats lie within the range $-0{\cdot}45<d/\rho<1{\cdot}8$ approx., and all the anastigmats within the range $-0{\cdot}36<d/\rho<1{\cdot}6$. (The anastigmat curve ends in a "singular point" at $R/\rho=1{\cdot}563\ldots$) Save in the trivial case of evanescence, none of these lenses fulfils the lateral achromatism condition of the " simple" eyepiece, and none is telescopic. Two flat-fielded aplanats exist which only need one surface figuring : one, figured on the front surface, corresponds to the intersection of curve A with the flat-field curve, having $d/\rho=-0{\cdot}41$, $R/\rho=0{\cdot}435$. To make this lens physically realizable, we must make ρ negative : it is thus a converging meniscus, presenting a concave face to the incident light. If we set $\rho=-1$, the focal length, F, is 2·559, and the astigmatic inter-focal distance is $0{\cdot}222\ldots h^2=0{\cdot}568\ldots h^2/F$, i. e. 0·568 times that of a centrally diaphragmed thinlens of the same focal length.

The second example, in which only the back surface is figured, corresponds to the intersection of curve B with the flat-field curve. It has $d/\rho=0{\cdot}98$, $R=1{\cdot}10$. If we set $\rho=+1$, $F=4{\cdot}75$, and the astigmatic interfocal distance is $0{\cdot}017\ h^2$, $=0{\cdot}081\ h^2/F$. The shape of this lens does not differ much from that of the $n=1{\cdot}618$ flat-fielded anastigmat described in the previous paper. If we "diaphragm" it at the pole of the back face, the first refraction will be nearly distortionless, while the second will provide the rather heavy pincushion distortion of fractional amount $\frac{1}{2}(n^2-1)\theta^2$ due to refraction from glass to air. The alternative design given above, supposed diaphragmed at the pole of the front face, will give the barrel distortion of fractional amount $\frac{1}{2}\left(\frac{n^2-1}{n^2}\right)\theta^2$ due to refraction from air to glass, since the distortion due to the second refraction will obviously be small. Both distortions are increased by shifting the stop to the point giving paraxial freedom from transverse chromatism, which lies a distance D' in front of the pole of the front surface, given by

$$D' = \frac{\rho d[(n-1)R - n\rho + (n-1)d]}{[n\rho-(n-1)d]^2 - R[n^2\rho-(n^2-1)d]}. \qquad \ldots\ldots (22)$$

(This expression was derived in the previous paper.)

Thus the $\rho=1$, $d=0{\cdot}98$ example has its lateral achromatizing point $0{\cdot}34\ldots\rho$ in front of the front surface, while the $\rho=-1$, $d=0{\cdot}41$ example has it $-0{\cdot}315$ in front of the front surface in object space—i. e. about 0·1 behind the back face in image space. When the lenses are "diaphragmed" at these points the respective fractional pincushion distortions at object angle θ are $0{\cdot}67\,\theta^2$ and $-0{\cdot}31\,\theta^2$.

It remains to calculate the strengths of the figurings needed in the two cases: in the artificial star-space the retardation required is m_1+m_2 (expression (2)). Transference to "fresh air" (i. e. air in front of the front face of the lens), since it increases the zonal height by a factor $\frac{n\rho}{n-1}$, multiplies plate strengths by $\left(\frac{n-1}{n\rho}\right)^4$. This gives for the figuring needed on the front face of the first $(\rho=-1)$ example a glass thickness of $3{\cdot}27\,h^4$ at zonal height h. Evaluating the corresponding quantity for the second example $(\rho=1)$ and mulitplying by $\left(\frac{L}{L-d}\right)^4$ to allow for the reduction of zonal height on passing to the back face, we find that the thickness required at height h on the back face is $0{\cdot}277\,h^4$. This completes the extended paraxial data for these designs.

§ 4. DISTORTION OF SINGLET APLANATS

I shall give, for the sake of completeness, expressions by which the distortion of these singlet aplanats can be computed. We consider, then, the plate diagram shown in figure 1 (in glass). It is convenient to divide the distortion into six contributions—three associable with refraction at the first surface and that part of the total figuring which aplanatizes it, and three associable in a corresponding way with the second surface.

If the principal ray make angle ω_s with the axis in a star-space of index n_s, and a plate of retardation mh^4 be situated at distance S down-light from the stop, the fractional pincushion distortion due to the plate is $\omega_s{}^2 4mS^3/n_s$. The aplanatizing figurings $\frac{m\alpha_\rho}{\alpha_\rho+\beta_\rho}$, $\frac{m\beta_\rho}{\alpha_\rho+\beta_\rho}$ defined in § 2 have together a third down-light moment about the ρ-centre of amount $m\alpha_\rho\beta_\rho(\alpha_\rho-\beta_\rho)$. So the fact that the refractions are aplanatized only, and not anastigmatized, provides, when the stop is located at the refracting centre, a pincushion distortion coefficient $\frac{4m\alpha_\rho\beta_\rho}{n_s}[\alpha_\rho-\beta_\rho]$, and correspondingly for the R-surface. (I retain the subscript s in n although for the glass star-space $n_s=n$, as a reminder of the origin of this factor.) Shifting the stop a distance S up-light increases the third moment of an aplanatic plate triad by $3S_s$ times its second moment. But the second moment was shown to be $m\alpha\beta$. Hence the *plate contribution* to the distortion is $\frac{4m\alpha\beta}{n_s}\cdot[\alpha-\beta+3S_s]$ with subscript ρ and R in turn. The remaining terms are the Petzval contributions which would operate if each refraction were anastigmatized by a plate at the refracting centre. If we set $S_s=S_{\rho s}$ for the case of the ρ-centre, and $S_{Rs}=S_{\rho s}+a$ for the R-centre, the Petzval contribution is

$$-\left(\frac{n-1}{2n}\right)n_s\left[\frac{S_{\rho s}}{\rho}-\frac{S_{Rs}}{R}\right]\omega_s^2. \qquad \ldots\ldots(23)$$

We find from (1)

$$a=\frac{R+d-\rho}{(\rho-L)(R+d-L)};\quad (\alpha-\beta)_\rho=\frac{(L-d)\rho+(\rho-d)L}{L(L-d)(\rho-L)};$$

$$(\alpha-\beta)_R=\frac{LR+(L-d)(R+d)}{L(L-d)(R+d-L)}, \qquad \ldots\ldots(24)$$

while $(m\alpha\beta)_\rho$ and $(m\alpha\beta)_R$ are the respective coefficients of H^4 in (4), so that the *plate contributions* are

$$\frac{-(n-1)n_s}{2L(L-d)}\left[(\rho-d)\frac{L}{\rho}\left(\frac{L}{\rho}-\frac{n+1}{n}\right)\right]\left[\frac{(L-d)\rho+(\rho-d)L}{L(L-d)(\rho-L)}+3S_{\rho s}\right]\omega_s^2$$

and

$$\frac{(n-1)n}{2L(L-d)}\left[(R+d)\frac{(L-d)}{R}\left(\frac{L-d}{R}-\frac{n+1}{n}\right)\right]\left[\frac{LR+(L-d)(R+d)}{L(L-d)(R+d-L)}+3S_{Rs}\right]\omega_s^2.$$

$$\ldots\ldots(25)$$

These formulae relate to the star-space of index $n_s=n$, defined in § 2. If we wish to transfer to a star-space of index 1, of the same focal length, we must delete the factor n_s and substitute for the brackets containing $S_{\rho s}$, S_{Rs} the values, n_s times smaller, of the corresponding distances in the new star-space. If we transfer to a star-space of q times greater focal-length, the corresponding distances will be increased q^2 times. When the object is at infinity it is natural to select as star-space "fresh air" in front of the lens. In that case put $n_s=1$ in the above; for a, $[\alpha-\beta]_\rho$ and $[\alpha-\beta]_R$, substitute $\frac{L^2a}{n}$; $\frac{L^2}{n}[\alpha-\beta]$; $\frac{L^2}{n}[\alpha-\beta]_R$; and for $S_{\rho s}$, S_{Rs} use the actual "fresh air" distances.

The distortion values given in § 3 were obtained in this way. But I have failed to find any method of simplifying (23) to (25), so the search for a distortionless flat-fielded singlet aplanat remains a problem of trial computations.

§ 3. REFLECTING FLAT-FIELDED MIRROR PAIRS

It is often salutary and sometimes instructive to set $n=-1$ in thick-lens formulae.

Let us set $\kappa=1$, $n=-1$ in (7) ; then the thick singlet becomes a mirror pair and the condition for flat-fieldedness reduces to

$$\frac{L}{\rho}\left[\frac{L}{\rho}(\rho-d)-\frac{L-d}{2}\right]=\frac{(L-d)}{R}\left[\frac{(L-d)}{R}(R+d)-\frac{L}{2}\right] \qquad \ldots\ldots(26)$$

or

$$\left(\frac{L}{L-d}\right)^2+\left(\frac{1}{2R}-\frac{1}{2\rho}\right)\frac{\rho^2}{(\rho-d)}\left(\frac{L}{L-d}\right)=\frac{\rho^2}{R^2}\frac{(R+d)}{(\rho-d)}. \qquad \ldots\ldots(27)$$

Let u be the distance of the object from the ρ-mirror, then

$$\frac{L-d}{L}=\frac{u(\rho-d)+d(\rho-u)}{u\rho} \qquad \ldots\ldots(28)$$

and

$$[u(\rho-d)+d\,(\rho-u)]^2+\frac{(R-\rho)}{2(R+d)}\,.\,Ru\,[u(\rho-d)+(\rho-u)]=\frac{R^2u^2(\rho-d)}{R+d}. \qquad \ldots.(29)$$

The corresponding conditions for anastigmatism instead of flat-fieldedness are obtained if in (27) or (29) the second term on the L.H.S.—the "Petzval"—is deleted.

Two object-positions are of special interest—at infinity, or on the surface of one of the mirrors. We shall first set $u=\infty$ when (29) reduces to

$$(\rho-2d)^2=\frac{R\rho[R+\rho-2d]}{2(R+d)}$$

or

$$R=\frac{\rho(-2d)}{2\rho}[\rho-4d\pm\sqrt{\rho^2+16d^2}]. \qquad \ldots\ldots(30)$$

It may be verified that this is identical with the condition given in another notation by Schwarzschild for extended paraxial flatness of the surface of best definition of an aplanatic two-mirror telescope objective. Such objectives have been very fully discussed by Schwarzschild, Chrétien and others. I shall therefore touch only on two points, about which I am frequently asked : Is there a design of two-mirror telescope O.G. in which (*a*) only the primary mirror, (*b*) only the secondary mirror should be asphierized ? The answer to both questions is yes, but the images in both cases are inconveniently placed—unless, indeed, one is prepared to figure the one figured mirror about a centre lying outside its rim, and so make an eccentric portion of an axially symmetrical objective—to make what might be called a Herschelian-Schwarzschild objective. When the primary mirror is to be figured, we must satisfy equation (14), which reduces to

$$\frac{d}{\rho}=\frac{R}{2d-\rho}\left[\frac{R+2d-\rho}{2R+2d-\rho}\right]^2. \qquad \ldots\ldots(31)$$

I think Schwarzschild himself pointed out the high "obstruction ratio" due to the secondary mirror in this case. Examples are $\rho=1$, $d=9/32$, $R=-7/32$ and $\rho=1$, $d=4/9$, $R=-1/9$.

When the secondary mirror is to be figured, we must satisfy (15), which reduces to

$$R=(\rho-2d)\left(\frac{\rho-2d}{\rho-d}\right)\left(1\pm\sqrt{\frac{d}{2\lambda-\rho}}\right). \qquad \ldots\ldots(32)$$

To make the primary mirror larger than the secondary, we must make ρ positive and $d<\rho$, while a real R demands $\rho<2d$. To obtain a real final image we have therefore to take the negative root. The final image of this rather peculiar "Gregorian" is awkwardly placed, lying between the primary image and the secondary. Example, $\rho=1$, $d=0{\cdot}7$, $R=-0{\cdot}1722$.

On the other hand, in the reflecting microscope ("projection" O.G.) we do not object to a secondary mirror several times bigger than the primary, and so we may set ρ negative in (32), with a physically real (positive) d, and take the negative root: examples, $\rho=-1$, $d=4$, $R=-5{\cdot}4$, $\rho=-1$, $d=7$, $R=-8{\cdot}91$. If it is desired to use apertures up to 0·65 N.A.—and this I find is possible—higher order aberrations cannot be neglected. (32) then gives only a rough approximation to optimal values, and the final spacing should be determined by ray tracing or—more simply—empirically. Expression (32) was given (in another form) by Schwarzschild.

Let us now consider the other special case, when the object lies on the surface of one of the mirrors. If we set $u=d$, i.e. put the object at the surface of the R-mirror, expression (29) reduces to

$$4(\rho-d)(R+d)=R\rho, \qquad \ldots\ldots(33)$$

the corresponding condition for anastigmatism being

$$4(\rho-d)(R+d)=R^2 \qquad \ldots\ldots(34)$$

But (33) is a physically symmetrical expression, for it states that the product of distances from the centre of each mirror to the surface of the other is equal to the product of the focal lengths. This suggests what may be verified formally, that since the object lies on the surface of one mirror, the image will lie on the surface of the other—an elegant and unexpected result. If, in addition to flatfieldedness, we demand anastigmatism, we must set $\rho=R$, giving $3\rho^2=4d^2$, and the magnification of the pair is $2+\sqrt{3}$.

REFERENCES

Burch, C. R., 1943. *Proc. Phys. Soc.* **55**, 433.
Martin, L. C., 1944. *Proc. Phys. Soc.* **56**, 104.

DEMONSTRATION:

A FRANCK AND HERTZ CRITICAL POTENTIAL EXPERIMENT

By J. H. SANDERS,
Clarendon Laboratory, Oxford

Demonstrated 8 *June* 1945

The experiment to be described is a repetition of the well known Franck and Hertz method of finding the critical potentials of mercury vapour (Franck and Hertz, 1914) suitable for lecture demonstration. It makes use of the cathode-ray oscillograph, and shows the use of the instrument, not only in demonstrations of this sort, but also in obtaining rapid qualitative observations in experiments involving two independent variables.

The electrode structure of the vapour-filled bulb is shown in figure 1. The electrodes are of nickel and the filament of 100-μ tungsten. The bulb was evacuated and the glass envelope, electrodes and filament were subjected to the usual out-gassing processes. A small quantity of mercury was distilled into the bulb, and this was then sealed off from the pumping system.

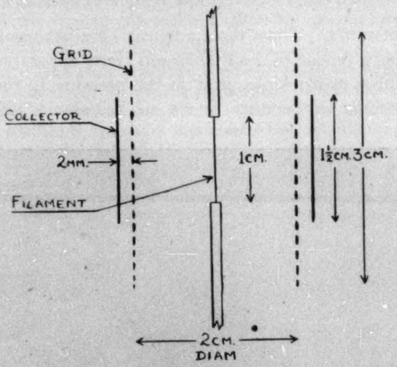


Figure 1. Electrode structure.

To show the curves relating collector current to grid voltage on the oscillograph screen, the oscillator shown in figure 2 is used. This delivers an alternating voltage of saw-tooth form between the filament and the grid. By using the cathode follower output, the grid swings from zero volts up to a positive voltage of about 30 volts peak with respect to the filament. The usual source of a small constant potential between the grid and collector is omitted, and the electron current flowing through the resistance in this circuit maintains the collector at a small negative potential with respect to the grid. The voltage across this resistance is amplified and applied to the Y-plates of the oscillograph, and the saw-tooth oscillator output is applied to the X-plates.

The shape of the curves obtained has been adequately explained elsewhere (Compton and Mohler, 1924). The shape depends greatly on the pressure of the mercury vapour in the bulb. The main kinks are due to the 4·7, 4·9 and 5·4-volt critical potentials. These three potentials are not resolved, though this has been done using a version of the Franck and Hertz apparatus (Morris, 1928). The size and number of the kinks grow as the pressure in the bulb is raised. The optimum working temperature seems to be about 120° c. A photograph taken from the oscillograph at this temperature is shown in figure 3. The first two kinks are separated by about 4·9 volts, and a smaller kink, due possibly to electrons undergoing one 4·9-volt and one 6·7-volt inelastic collision, can be seen on the side of the second main kink. The kinks appearing at higher grid voltages are due to electrons undergoing various combinations of the 4·9-volt

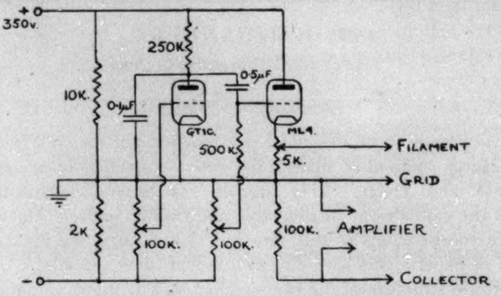


Figure 2. 50-cycle saw-tooth oscillator.

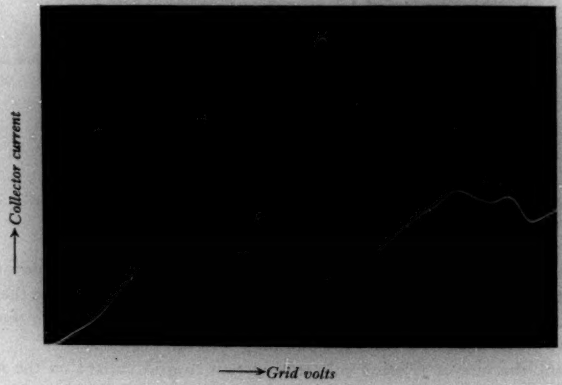


Figure 3. Curve taken at 120° c.

group of collisions and 6·7-volt collisions, as explained in the paper of Compton and Mohler.

The change of shape of the curves as the vapour pressure of the mercury in the bulb is altered may be demonstrated in a few minutes by using the oscillograph, this representing a very considerable advantage over the more conventional " D.C. " approach to the experiment.

I am indebted to Mr. C. H. Collie and Dr. D. Roaf for their help in setting up this experiment.

REFERENCES

COMPTON and MOHLER, 1924. *Bull. Nat. Res. Council, Wash.*, no. 48.
FRANCK and HERTZ, 1914. *Verh. dtsch. Phys. Ges.* **16**, 10.
MORRIS, 1928. *Phys. Rev.* **32**, 447.

DISCUSSION

Dr. T. B. RYMER. I should like to ask Mr. Sanders whether it is possible to use any type of commercial thyratron valve for a demonstration of this kind, in place of the specially constructed valve we have seen.

Mr. A. E. JENNINGS. Assuming that the author derived his Y deflection, which was a measure of the current passing to the outer electrode, by applying the potential drop across a high resistance in series with the small constant retarding potential to the Y-plates of the C.R.O., did this varying potential not affect the constant retarding potential, and how did this affect the validity of the experiment?

AUTHOR'S REPLY. The most important condition which has to be fufilled in order to obtain large sharp kinks is that the drift space between the filament and the grid shall be large compared with the distance between the grid and the collecting electrode. This condition was not fulfilled by the only thyratrons which were available, and it was therefore thought better to make up a suitable valve.

Mr. Jennings is of course right in pointing out that the collector potential is not constant in this experiment. The effect of the 100 K. resistance between collector and earth will be slightly to diminish the sharpness of the kinks, as compared with the curves taken with a fixed retarding potential. As is seen in figure 3, this effect is not serious.

DEMONSTRATION:

APPARATUS ILLUSTRATING THE MICHELSON STELLAR INTERFEROMETER

BY P. J. TREANOR, S.J.,
Clarendon Laboratory, Oxford

Demonstrated 8 *June*, 1945

THE effect upon which this instrument operates is the disappearance of the interference fringes formed by two similar apertures by the appropriate displacement of the patterns arising from neighbouring sources of light (Michelson, 1890, 1920). In a laboratory experiment the relation between the angular separation of the objects to be resolved and the separation of the diffracting apertures is most conveniently altered by varying the former, since a double slit of variable separation is rarely available. The use of a Fresnel bi-prism, and like devices for producing artificial doubles, is prohibited by the fact that the components so produced are coherent sources, and do not therefore simulate a genuine double star.

These difficulties are overcome by the use of a Wollaston quartz compound prism, which is placed (figure 1) on an optical bench in front of an illuminated pin-hole. The prism produces the components by double refraction, the light from each being polarized in mutually perpendicular directions. Since the components are selected from different planes of vibration of the natural light from the pin-hole source, they are genuinely incoherent, in the sense that they would not interfere, even if brought back into one plane by introducing a polaroid plate.

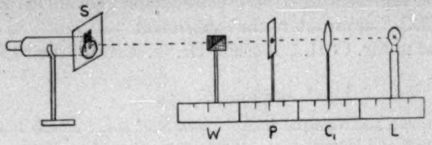


Figure 1.

The effect is observed through a small telescope, whose objective is covered by an appropriately oriented double slit. It is convenient to cover part of the slit with a weak cylindrical lens, which has the effect of producing a greatly broadened pattern, without altering the distance between successive maxima. The telescope should be a couple of metres from the pin-hole. Slits 2 mm. apart, of width 0·3 mm., with a cylindrical lens of 0·25 dioptres, give easily observed fringes in a telescope of 30 cm. focal length, × 10. Their disappearance may be observed for various separations of the components, produced by moving the Wollaston along the bench.

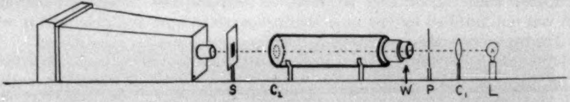


Figure 2.

Figure 2 indicates how the artificial double may be used in a very similar way when it is desired to photograph this and other diffraction phenomena connected with double-star resolution. The apparatus is here adapted to the optical system of a spectrograph of 150-cm. focal length, the collimating lens being the only important modification to the arrangement of figure 1. Owing to the incoherence of the light from the components and their different planes of vibration, a polaroid may be introduced between collimator and camera to vary the relative intensities of the components. I have used the diffraction photography arrangement of figure 2 to photograph the advantageous effect which a square stop affords in resolving artificial doubles of very unequal intensity.

REFERENCES

Couder and Jacquinot, 1939. *C.R. Acad. Sci. Paris,* **208,** 1639.

Michelson, 1890. *Phil. Mag.* **30,** 1 ; 1920. *Astrophys. J.* **51,** 257 ; 1921. *Ibid.* **53,** 249.

OBITUARY NOTICES

JOHN AMBROSE FLEMING

WITH the death of Sir Ambrose Fleming the Physical Society loses the last survivor of its Founder Fellows. That bald statement does less than justice to a connection which is unique in the history of scientific societies, for on 21 March 1874, at the inaugural meeting of the Physical Society, the first communication to be presented to the Society was a paper by J. A. Fleming on the theory of the galvanic cell. During the long and distinguished career that followed this auspicious opening, Fleming contributed more than one hundred important papers to the Proceedings of various learned societies and institutions, and of these papers more than a third have appeared in the *Proceedings of the Physical Society*. The interval between the first and the last of these communications was no less than sixty-five years, for on 28 April 1939 Sir Ambrose Fleming read before the Society a paper on " A new method of creating electrification ", illustrated by experiments. The occasion was one never to be forgotten by those who were privileged to be present. The veteran author was in fine form, and his presentation of his paper, made, as it was, in a clear and resonant voice, not read from a manuscript, logically developed without the introduction of a superfluous word, was a model for speakers not a quarter of his age. Fleming, early in his career, had obviously profited by the classic advice given to young debaters : " State your case, clear your case, prove your case : sit down "; his published works bear eloquent testimony thereto.

John Ambrose Fleming was born on 29 November 1849 at Lancaster, where his father was a Congregational Minister. When he was about five years old, his parents removed to North London, where his youth was spent. Quite early in his life, he, like Clerk Maxwell before him, showed a decided inclination to investigate the " particular go " of any mechanism that came his way, and he equally early developed, with home-made apparatus, that technique of popular lecturing in which, in the years of his maturity, he was to prove himself a worthy successor to Faraday and to Tyndall.

After some years of attendance at private schools he was sent, at the age of fourteen, to University College School, then housed in a wing of University College in Gower Street. He had (in common with many other lads of similar tastes) ambitions towards the engineering profession, but the usual mode of entry thereto—a long apprenticeship and a high premium—was out of the question, so it was decided that he should work at school for the London Matriculation Examination, which he passed at the age of sixteen. Then followed two years at University College, where he studied under Carey Foster, A. W. Williamson and—an influence for life—that great teacher and erudite genius Augustus de Morgan. At eighteen he passed the first B.Sc. examination of London University, but the question of earning a living loomed large and, after a fruitless and unhappy three months in the offices of a shipbuilding firm near Dublin, he returned to London to take up work as a clerk on the Stock Exchange where, the hours of work (10 a.m. to 4 p.m.) being short, " I had a good deal of time for private study ". It was a hard task he had set himself but he won through and the 1870 pass list for the degree of B.Sc. in the University of London contains the name of John Ambrose Fleming in the first division.

He now bade farewell to the Stock Exchange, though he never regretted the time he had spent there, for the knowledge of business methods thus acquired proved very useful in his later consulting work. Through the good offices of Frankland the chemist, he obtained the post of Science Master at Rossall, a public school on the Fylde coast between Blackpool and Fleetwood. Here he taught with characteristic energy and success for a year and a half and, equally characteristically, determined that he had " in some senses more need to be still a student rather than a teacher ". He resigned accordingly in the autumn of 1872 and immediately entered Frankland's laboratory at the Royal College of Science in Exhibition Road, where at that time Frankland professed chemistry, Guthrie physics and Huxley biology. He became research assistant to Frankland and took a share in the teaching in the department. It was during this period that he read the paper, previously mentioned, at the inaugural meeting of the Physical Society.

In 1874 he was appointed science master at Cheltenham College on the civil and military side and was now embarked on what for many would be a life's work; but Fleming was by nature a student and an engineer. While at Cheltenham he read with care Faraday's *Experimental Researches* and, possessing a fairly strong electromagnet, repeated on a small scale Faraday's abortive experiment to detect electric currents induced in a river flowing across the earth's magnetic field. Fleming sent a stream of acidulated water down a tube placed between the poles of the magnet and carrying, fixed to opposite sides by wires sealed through the tube, two strips of platinum foil. When the magnet was energized, the strips exhibited the characteristic polarization effect. Fleming described this result in a paper read before the British Association at its Bristol meeting in 1875.

But Fleming had also been reading Maxwell's recently published treatise, and, coming to the conclusion that his mathematical equipment was defective, resigned from his safe £400 a year post and entered Cambridge in October 1877 as an Exhibitioner of St. John's College.

One of Fleming's principal objects in going to Cambridge was to study and work under Maxwell, who assigned to him the task of comparing the various coils that had been constructed to represent the B.A. unit of electrical resistance and to deduce its most probable value. To that end Fleming constructed a special form of Carey Foster bridge, promptly christened by Maxwell " Fleming's banjo ". He took careful and copious notes of Maxwell's lectures, and presented the bound volumes of notes to the Cavendish Laboratory on the occasion of the celebration of the centenary, in 1931, of Maxwell's birth.

In 1879 Fleming took his D.Sc. degree in London and in 1880 graduated in Cambridge by way of the Natural Science Tripos. In the same year he was appointed Demonstrator in Mechanism under Professor James Stuart and in 1882 was elected to a Fellowship at St. John's College. In the meantime he had left Cambridge, having been appointed in 1881, through the influence of Professor Stuart, to the chair of Mathematics and Physics at University College Nottingham.

But events were impending which were to make his tenure of the Nottingham post of short duration. It is difficult for those of us who saw and heard the veteran in full vigour so lately as 1938 to realize the magnitude of the changes, social and scientific, which had occurred during his lifetime and which he himself had helped so largely to shape. In the year of his birth Peel, Palmerston, Russell and Derby were outstanding political figures—Gladstone and Disraeli were, politically, promising youngsters ; in the practice of war, Wellington and Nelson could have handled the armies and navies of the period without recourse to much new learning ; the stage-coach had not yet vanished from the scene ; the electric telegraph was a novelty.

In 1881 the storage cell was just about to become a practical proposition ; Bell had invented the telephone and, realizing that the Bell telephone was an excellent receiver but a poor transmitter, Edison had invented his carbon-button transmitter ; the possibilities of alternating currents were beginning to loom large, and practice in that branch of science, running ahead of theory, was to provide in the " eighties " some signal shocks and surprises for station engineers ; Edison and Swan, in attempts to realize electric lighting on a commercial scale, were performing Heath-Robinsonesque experiments to produce filaments of carbon ; switches, mains, cables and measuring instruments were urgently needed ; patents of almost every type were filed ; patent litigation, with its attendant demands for scientific witnesses (how should one, for example, define exactly a filament of carbon ?) were frequent ; and Fleming, who had been appointed to act as a scientific adviser to the Edison Telephone Company in 1879, was appointed as electrician to the Edison Electric Company in 1882.

In 1885 Fleming was invited to occupy the newly established Chair of Electrical Engineering at University College London. The only equipment attached to the Chair at the time of the appointment was chalk and a blackboard, but a timely grant of £150 provided funds for a gas-engine, a battery of storage cells, a photometer and some essential electrical measuring instruments. Such was the modest beginning of the Pender Chair of Electrical Engineering. Fleming resigned in 1926, after forty-one years tenure of the Chair. During that period he had made the Pender Laboratory world-famous ; upwards of two thousand students had passed through his hands, many of whom achieved great distinction in after-life ; a number of highly important papers were published by the Pender Professor, including an early paper read before the Institution of Electrical

Engineers in 1885 on the necessity for a national standardizing institution for electrical instruments—a paper which foreshadowed the establishing of that great organization, the National Physical Laboratory. Later came the results of measurements made on alternating-current transformers which cleared away many difficulties, pioneering researches on electric-lamp photometry, and a classic series of papers published in collaboration with Sir James Dewar on electrical properties of matter at very low temperatures.

All this research and teaching was diversified by a large amount of consulting work, the most important of which arose from his appointment as consultant to the Marconi Wireless Telegraph Company in 1899.

At that time wireless telegraphy was rapidly approaching the commercial stage ; Marconi had transmitted messages in Morse between the Isle of Wight and Bournemouth ; and in September 1899, at the meeting of the British Association at Dover, Fleming gave a brilliant evening address at which wireless messages were exchanged between the British Association and the French Association then meeting at Boulogne. Obviously, as with the corresponding developments in the electric telegraph half a century earlier, men's thoughts were turned to the possibility of uniting the Eastern and Western hemispheres, and one of Fleming's early tasks in his new position was the designing and equipping of a transmitting station for that purpose at Poldhu in Cornwall. The power required must have been assessed by the exercise of what the Greeks termed "happy guessing", but however that may be, a twenty-five h.p. engine, and an alternator and transformer capable of supplying current at twenty thousand volts, were installed ; aerials, switches, jiggers and condensers were designed on an engineering scale ; and, despite many difficulties and delays, in which stormy weather played no small part, the station was put into working order. In November 1901 Marconi sailed for Newfoundland, and on 12 December of that year—a red-letter day in the history of electrical science—he received the signals S-S-S - - - - transmitted from Poldhu.

Important as was Fleming's share in this triumph, his outstanding contribution to wireless telegraphy was to be made some three years later. As far back as August 1885 he had read a paper before the Physical Society in which he had noted the presence of a clear white line, or line of no deposit, on the inside of the bulb of a blackening carbon filament lamp and in the plane of the filament. He called this a molecular shadow, and gave the obvious interpretation of the phenomenon. In the previous year Edison had observed the "Edison effect" in a carbon filament lamp. He had fixed a metal plate between the hairpin bend of the filament of a lamp and had discovered that when the filament was heated by a current, a galvanoscope attached between one terminal of the filament and the plate indicated a slight current ; when the attachment was made between the plate and the other terminal of the filament no current was observed. Edison gave no explanation of the phenomenon and made no notable application of it. Fleming carried out an elaborate series of experiments on cylinders and plates of various kinds sealed into filament lamps, and concluded that the heated filament was emitting negatively electrified particles, which he assumed to be carbon atoms, and that the space between filament and cylinder was a one-way street for electricity.

In the early years of the twentieth century the problem of the production of a trustworthy rectifier for feeble high-frequency oscillations was an urgent one, and in October 1904 it struck Fleming that his old experiments had provided a solution to the problem. Investigation showed that this was indeed true, and on 16 November 1904 he applied for a British patent.

Such, in very brief outline, was the origin of the Fleming valve.

Fleming had unusually high gifts as a lecturer, and his services were in constant demand over a long period of years by various educational institutions. He served for many years as a lecturer under the Gilchrist Educational Trust ; he delivered eight Friday evening lectures at the Royal Institution, and between 1894 and 1922 he gave the Christmas lectures at the Royal Institution four times ; his public lectures were a feature of many of the annual meetings of the British Association ; and he gave the Cantor series of lectures for the Royal Society of Arts no less than eight times. His last public lecture was given before that Society in 1925 on "William Sturgeon and the Electromagnet".

His books, too, exerted a profound influence on the teaching of the period, and thousands of students of many generations have read and re-read his treatises on the transformer and on wireless telegraphy. It is a notable fact that the veteran occupied the last years of his long life in composing a series of four text-books for the use of students of engineering.

It is surely a feat unique in the history of education that a nonagenarian should produce books which should be useful to students of these later-day developments in physical science. True, there are a few slips here and there, but the old Fleming touches are still to be discerned: terseness and clarity, not a wasted word, and an essentially practical outlook.

Many and well-deserved honours fell to his lot. In 1892 he was elected Fellow of the Royal Society, and he was awarded the Hughes Medal of the Society in 1910. In 1921 he was awarded the Albert Medal of the Royal Society of Arts. The Institution of Electrical Engineers gave him their Faraday Medal in 1928. In 1930 the Physical Society awarded him the Duddell Medal. In 1935 he received the Kelvin Medal from the Institution of Civil Engineers, and the Franklin Medal of the Franklin Institute. In 1929 he received the honour of knighthood. After his retirement from teaching he lived at Sidmouth in Devonshire, and in 1933 he married Miss Olive Franks. Active and interested in outside affairs to the last, he died on 19 April 1945.

There is nothing for tears, nothing for sorrow, in the passing of a life so honoured, so prolonged, so full of work which has left its mark on the scientific and social life of our era; yet, in taking an affectionate leave of the last survivor of their Founder Fellows, the Physical Society cannot help but regret that he was not spared to see the centenary of his birth and the 75th anniversary of the day on which he inaugurated the work of the Society by reading before it the first paper in their *Proceedings*.

A. F.

SIDNEY SKINNER

Sidney Skinner (23 August 1863 to 5 November 1944) was educated at Dulwich College, whence he proceeded to Christ's College, Cambridge, with a scholarship in 1882. In 1884 he took the Natural Science Tripos, obtaining a second class in Part I, and a First in Part II in 1886. He stayed on at Cambridge, being appointed a lecturer both at the Cavendish Laboratory and at Clare College. He was also an assistant demonstrator at the Laboratory. In 1890 he was appointed Demonstrator, assisting Wilberforce (afterwards Professor at Liverpool) with the senior students. Searle was appointed a demonstrator at the same time and had charge of the first-year students. Other demonstrators with him were Fitzpatrick (afterwards President of Queens') and Glazebrook (later Sir Richard).

In 1904 he became Principal of the South-Western Polytechnic at Chelsea, London. At that time this Polytechnic—as all the London ones—taught a large variety of subjects; but it gradually became more specialized, and under the influence of Skinner it concentrated upon preparation for the Internal Degrees of the London University in Science, being especially strong on the biological side. The only other activities of the Polytechnic were a school of pharmacy, a good Art School and a school for preparing women to become teachers of physical training.

In 1906 he went with the expedition to Spain to see the eclipse, and there met and married Marian Field Michaelis, who had also gone there from Canada. She was daughter of Major Michaelis. Mrs. Skinner was a very live woman who took a keen interest in science; both she and her husband were regular attendants at the Friday evening lectures at the Royal Institution, of which Institution he was for some years a Visitor.

They had three children, the youngest a son. Both Mrs. Skinner and this son are dead.

On his retirement from the Polytechnic, in 1928, he went to live at Letchworth. There he collected a group of enthusiasts who held informal meetings and discussions on scientific subjects and were facetiously referred to by their friends as " the Wise Men ".

He was a Fellow of the Institute of Physics, of the Physical Society and of the Cambridge Philosophical Society. He published nearly forty papers on a variety of subjects, the most important being on *The Clark Cell as a Standard of E.M.F.*, on *Lubrication*, and on *Dew, Dew-ponds and Humidity*.

His greatest interest was in climbing. From 1892 to 1907 he spent nearly every summer vacation in Switzerland. Two of his papers were connected with this recreation, *On the Slipperiness of Ice* and on the *Minute Structure of the Surface Ice of Glaciers*.

Skinner was a quiet conscientious worker who never cared to be in the limelight, but he had a wide circle of very real friends who appreciated his genuine ability and kindliness.

R. S. C.

LIONEL ROBERT WILBERFORCE

Born at Munich on 18 April 1861, Lionel Robert Wilberforce died on 1 April 1944 in his 83rd year. He was a direct descendant of William Wilberforce, pioneer in the abolition of the slave trade.

He was elected a life Fellow of the Physical Society in 1894 and for a period he served on the Council—his membership of the Society covered a period of almost exactly 50 years.

Educated at the London International College, Isleworth, and Trinity College, Cambridge, he obtained a first-class honours in the Natural Science Tripos and was 13th wrangler. He served as demonstrator and lecturer in Physics under J. J. Thomson from 1887 to 1900, when he was appointed Lyon Jones Professor of Physics at Liverpool University in succession to Sir Oliver Lodge. He was almost entirely responsible for the planning of the George Holt Physics Laboratories and took a great pride in the work.

Wilberforce was well grounded in the so-called " classical " physics, but took little more than a tolerant interest in the " modern " physics—quantum theory and atomic physics. As a teacher of physics he was unique. His lectures were illustrated by a plentiful display of experiments which he had specially designed and carefully rehearsed. His obvious enjoyment in demonstrating a point by a successful, and in many cases original, experiment spread to his audience.

One of his favourite experimental lectures was that relating to the principles of the gyroscope. This was always a great success and was very popular. His running commentary whilst manipulating a " diabolo " is not likely to be forgotten by his audiences.

His opinions on fundamental classical physics were sound and reliable, but he fought shy of expressing views on " modern " atomic physics. His work on surface tension, viscosity and allied problems is well known from his publications, but much of his work on these and other subjects was not published. He developed a beautiful ripple tank to demonstrate wave phenomena such as diffraction, reflection, refraction, interference, etc. I remember seeing this for the first time in 1913 and was much impressed by its possibilities in demonstrating acoustic and optical phenomena. His keen interest in instrument design is illustrated in his paper " Kinematic supports and clamps ", read before the Physical Society in November 1932.

There is no doubt, however, that he will best be remembered by his students as a great teacher of physics.

He appreciated a good joke and was fond of relating his experiences in examining students. I remember his enjoyment in telling of a student he was examining in elementary practical physics at Cambridge. The subject was the spirit-level. The bubble moved to the right, the student concluding that the right-hand end of the table was higher than the left. On reversing the spirit-level, the bubble moved to the left. The student looked suspiciously at the spirit-level, then at Wilberforce, and said, after a pause, " Why, the bally thing's cockeye ", a correct, if somewhat unscientific, conclusion.

Wilberforce was a member of the Alpine Club and in his younger days did a good deal of climbing in the Alps. He was very fond of skating on ice and on rollers. Such vigorous exercise, which he kept up until his later years, no doubt contributed to his rosy complexion and general good health.

It was a great pleasure to have known him. His pleasant disposition, sound judgment and success as a teacher of physics are likely to be remembered by all his many students and friends.

A. B. WOOD.

JAMES YOUNG

James Young, who died at the age of 47 in May 1945, after a year's unflinching fight against leukaemia, had been a physics lecturer in Birmingham University since 1920. A Scot from Strathaven, his student years had been spent in London at the Royal College of Science, and when Dr. S. W. J. Smith, F.R.S., left that College to succeed Poynting in the Birmingham professorship, he shrewdly selected Young as one of the two recent graduates of outstanding promise to accompany him.

Young possessed a combination of gifts rare in contemporary physics. He took aesthetic delight in precision, both of numerical accuracy and of logical proof, and regarded with tolerant astonishment anyone's liking for the stages in a subject which only yield rough exploration or ambiguous conclusion. Nevertheless he was outstandingly successful at teaching elementary students whose preoccupations would seldom dispose them to any comparable rigour ; for he had the personality to command respect from those who could not approach his own intellectual ideals, and he was devoid of arrogance or contempt. He would, in fact, give limitless care to any honest enquirer, student or colleague or stranger.

This quality of the intellectual aristocrat often made Young seem lonely among any group who were advancing more rapidly in subjects less susceptible of certitude or minute precision, and he was not a very communicative colleague. Perhaps few beyond the writer heard him explain or confess that his real enthusiasm was for the geometrical in physics. This constituted him a referee of unrivalled critical judgment, rather than an explorer of new grounds. For instance, he employed his genius for three-dimensional visual imagery in elucidating complex crystalline structures, and his X-ray contributions to Prof. S. W. J. Smith's researches on nickel iron and to Prof. W. N. Haworth's new carbohydrates were notable : the same skill made him in war-time the expert on polar diagrams for Prof. M. L. Oliphant's team working at radar projection. But he had no taste for pushing into the crowd who throng the more sensational avenues to scientific eminence, and he probably spent far more time on measuring photographs of lunar craters or extracting the final decimal place in timing an occultation than on any more obviously rewarding topic.

In advanced expositions, including very individual and physical treatment of higher mechanics, trigonometrical astronomy, and Lorentz-Einstein relativity, he bequeathed to students something of his rare gifts for seeing geometry in physics : their notebooks may well have a clarity unsurpassed in all the library of treatises which Young disdained to copy and was too self-effacing to rival in print. MARTIN JOHNSON.

REVIEWS OF BOOKS

Practical Optics, by B. K. JOHNSON. Pp. viii + 189. (London: Hatton Press, 1945.) 15*s*. net.

The greater precision that has been increasingly asked for in the engineering industry has resulted in the introduction of new instruments both for the measurement of dimensions to finer limits and for the closer examination of the finish. These instruments have largely made use of optical elements and devices. Thus, in addition to the older users of optical instruments—microscopists, photographers and astronomers—there has now been added a new type of user, the engineer. All these users, both old and new, want to know enough about the optics of these instruments to be able to select the instrument that will best suit their purpose and to enable them to use it intelligently, but at the same time they do not want to spend time learning more of optics than they are obliged to do, for their main interest is elsewhere. For them, therefore, a book that will just give them this minimum knowledge is a valuable time-saver. Such a book is the one under review.

It begins with some simple demonstrations which such a user should, if possible, repeat for himself, as he will find that he learns far more easily how his instruments behave from these simple experiments than he can do from merely reading about them. The principles upon which the several instruments depend are clearly explained, in every case being closely related to practical applications.

Owing to the aberrations which are present in all optical systems, every instrument is necessarily a compromise, and sacrifices some features in order to attain the greatest excellence in those which are of the greatest importance for its particular purpose. The telescope, for instance, in order to obtain a very high definition in the centre of the field, is content with an angular field which, in comparison with that of a photographic lens, is exceedingly small, and with a numerical aperture which in comparison with a micrographic object is again exceedingly small. Thus, before selecting an optical system for any particular purpose—especially if this is a new departure—it is necessary to know what are the special requirements that the system must satisfy, and the explanations given in this book will greatly assist in this selection.

We have noted the following printer's errors. The caption to figure 90 *a* should read " Measurement of Numerical Aperture ", and that of figure 127 should read " Cooke Anastigmat ". In the table at the foot of page 109, the last column should read 38·6, 37·0, 35·2. R. S. C.

Théorie des Oscillateurs, by YVES ROCARD. Pp. 220. (Paris: Editions de la Revue Scientifique, 1941). Price not stated.

M. Rocard believes that too many students of physics think that all oscillations are fundamentally similar to those of a simple pendulum, and he has written a pleasant monograph to correct this over-simplified view. The whole tone of the book is critical and, if one may be allowed to say so, refreshingly old-fashioned ; old-fashioned in the sense that it is obviously written to tell the reader about oscillations and not, as are so many books on mathematical physics, written with the main object of introducing the reader to new mathematical methods.

The first chapters of the book are a critical introduction and commentary on the series of papers published by van der Pohl in the *Philosophical Magazine* on relaxation oscillators. The treatment given is elementary and straightforward and has the merit of working specific problems through in detail. In particular, the analysis of coupled oscillators, in which the causes of frequency hysteresis are explained, and the conditions for the locking of frequency-multiplying or dividing circuits obtained, will be helpful to those who have to cope with the vagaries of multivibrators. Equations such as

$$\ddot{r}-2\omega\left[1-r^2-\frac{\dot{r}^2}{\omega^2}\right]\dot{r}+\omega^2 r=0,$$

which have strictly harmonic solutions, are used to poke gentle fun at some of the qualitative generalizations which are sometimes used in explaining the action of this type of oscillator. Indeed, throughout the book the tendency is to consider specific differential equations which can readily be seen to govern the motion of a simple (usually mechanical) system, and to avoid generalities.

In this way the reader is naturally introduced to systems governed by equations of the type

$$a_0\frac{d^n y}{dt^n}+a_1\frac{d^{n-1}y}{dt^{n-1}}\cdots+a_n y=0.$$

Routh's criterion for the instability of the system are introduced and their application explained in several examples based on mechanical models. This is an excellent chapter, and much ingenuity has been exercised in devising the rather artificial models.

More natural examples are provided in the following chapter, which deals with the differential equations which have been used in an attempt to analyse the economic oscillations of the trade cycle. The author is somewhat critical of this approach on the ground, which he justifies by numerous examples, that the predicted oscillations are unduly sensitive to small changes in the ideal mechanisms used to represent the complexities of economic life. Kalecki's equation and others which have been proposed are characterized by a period of gestation or lag between economic cause and industrial effect.

M. Rocard makes a positive contribution to the subject by pointing out that not only is the assumption of an exact value τ for the time lag a very artificial one, but also one which greatly enhances the difficulty of solving the equations. The much more plausible assumption that a lag τ has a probability given by some such function as

$$\left(e\frac{\tau}{\tau_0}\right)^n e^{-n\tau/\tau_0}$$

gives rise to equations which are much more tractable.

For example, a simple relation of the type

$$y_{(t-\tau)}+C\dot{y}=0$$

based upon a definite period of gestation τ becomes (for $n=1$)

$$A\int_0^\infty \tau e^{-\tau/\tau_0}y_{(t-\tau)}\,d\tau+\dot{y}=0,$$

which readily reduces by the substitution $y=e^{pt}$ to the form

$$p^3+2\gamma p^2+\gamma^2 p+A=0,$$

to which Routh's criteria can be applied directly.

The remaining chapters deal with waves on lines, the hunting of coupled generators, organ pipes, and the flutter of aeroplane wings.

While these subjects are interesting in themselves, they are less suitable for a purely mathematical treatment, and do not add to the refreshing Gallic unity of the previous chapters. The section on lines, while adequate for 1940, is not of great interest to present-day students. The book was written during the fall of France, and the author asks the indulgence of his readers for the absence of references. The typography is good and the paper better than British readers have seen for some time. The absence of an index, which is partly compensated for by a full table of contents, shows that insularity occurs in the Ile de la Cité as well as at home. On the whole a book which the physicist with a taste for elementary mathematics would do well to buy.

C. H. C.

Life and Work of John Tyndall, by Professor A. S. Eve, F.R.S., and C. H. Creasey. Pp. xxxii + 404. (London: Macmillan and Co., Ltd., 1945.) 21*s.* net.

We have in this book by Professor Eve and Mr. C. H. Creasey a fairly complete story of the life of John Tyndall, from his boyhood at Leighlin Bridge, Co. Carlow, Ireland, where he was born on 2 August 1820, to his death on 4 December 1893. The development of the poor country lad into the head of the Royal Institution is an interesting human story. Tyndall was fortunate in that the master of the National School at Ballinabranagh, which he attended as a youth, was capable of giving him a sound training in mathematics and the rudiments of physical science, as well as in the English language.

Tyndall's first appointment was in the Ordnance Survey Office at Youghall, Co. Cork, and later he was transferred to Preston, being dismissed from the service after three and a half years for joining with 18 other assistants in an appeal against their conditions of work. After an interlude as a surveyor on new railway schemes, Tyndall obtained in 1847 a post as a master at Queenwood College, where he had Edward Frankland as a colleague. A year later Tyndall and Frankland left Queenwood together to go to the University of Marburg, where Tyndall graduated as Ph.D. after two years. While at Marburg he studied under Bunsen. One of Bunsen's assistants was Dr. Heinrich Debus, later Professor of Chemistry at the Royal Naval College, Greenwich, and one of Tyndall's closest friends. Professor Knoblauch had then just arrived at Marburg, from Berlin, and from him Tyndall got his first stimulus to study magnetism.

On his return to England he found some difficulty in obtaining a congenial post. He and T. H. Huxley were unsuccessful in their applications for professorships at the University of Toronto in the summer of 1851 and at the University of Sydney a few months later. During this rather anxious period Tyndall was actively engaged in various researches, and was elected Fellow of the Royal Society in June 1852. In the following year he was appointed Professor of Natural Philosophy at the Royal Institution, where Faraday was Resident Professor. Faraday and Tyndall remained always on terms of the closest friendship in spite of the fact that they held different opinions on a number of scientific questions. On Faraday's death in 1867, Tyndall succeeded him as Resident Professor in the Royal Institution, and it was at the Royal Institution that Tyndall's active scientific life was passed. From the autumn of 1859 until the summer of 1866, Tyndall also held the Professorship of Physics at the School of Mines.

Tyndall was an eminently successful lecturer, and the books in which he published the content of his series of lectures are still very readable today, and have inspired many a young man with the desire to follow a scientific career. He devoted the greatest care to devising good experiments to illustrate his lectures, and the technique of the physical laboratory owed much to his ingenuity. His best known experimental investigations were those in diamagnetism, radiation and absorption by gases, and his very thorough demonstration that the moulds which Bastian had found to originate in infusions of hay, meat, fish, etc., were due to germs in the atmosphere, and were only produced when the infusions were exposed to the atmosphere. Tyndall succeeded in completely demolishing Bastian's ideas concerning the spontaneous generation of life. He also showed that the blue colour of the sky is due to scattering of light, though he ascribed the scattering to small particles and droplets suspended in the atmosphere, and it was left to the late Lord Rayleigh to demonstrate that it is the molecules of air themselves which are effective in producing the blue colour.

Tyndall derived from his father a tendency to religious emotion and controversy, and all his life he was apt to be drawn into controversies on scientific topics of all kind. He was in no sense lacking in courage, and had no hesitation in expressing his opinions even when they were unpopular. All his life he was fascinated by glaciers, and his studies of the motion of glaciers, which incidentally led him into controversy with J. D. Forbes, were among the earliest serious scientific studies of Alpine phenomena. Tyndall also devoted a great deal of energy to championing the claims for priority of scientific men whom he regarded as neglected. Thus he strongly championed Mayer's claim to have first put forward the principle of the conservation of energy.

From this book the reader receives an impression that controversy filled quite a large part of Tyndall's life. But at the same time it is made clear that Tyndall had a gift for warm friendship with men of the widest variety of interest—Faraday, T. H. Huxley, Charles Darwin, T. A. Hirst, Thomas Carlyle, Alfred Tennyson, Hooker, Lubbock. Tyndall must, in fact, have been a great social success.

It cannot be claimed that Tyndall had the spark of divine fire which Faraday had, but he was a sound and skilful experimenter, with a marked gift for exposition. He had no outstanding discovery to his credit, but he discovered many new facts, and explained many things which had hitherto been obscure. His books, which are very largely based on lectures given at the Royal Institution, formed a considerable contribution to education in the science of physics. His experimental technique was frequently far in advance of that of his contemporaries, and his breadth of interest brought him into touch with many of the vital problems of his day.

This biography of Tyndall is an important contribution to scientific history. It has a somewhat uneven style, partly due to the fact that a first draft was completed by the late Professor A. S. Eve just before his death, and was revised by Mr. Creasey. Material for an authoritative biography of John Tyndall had been collected by his wife, who survived him by nearly 47 years. Mrs. Tyndall had from time to time endeavoured herself to write the biography, and had at first endeavoured to control any effort made by others to complete the work. The unevenness of style is to be explained by the vicissitudes of authorship. It need not, however, deter anyone interested, either in the science of physics or in its history, from reading the story of a full and busy life, lived at the centre of English scientific activity, among those giants of the Victorian age—Faraday, Kelvin, Maxwell, Rayleigh, Stokes—whose work was to revitalize science and to lay the solid foundation on which later generations have built. **D. B.**

An Introduction to X-ray Metallurgy, by A. Taylor. Pp. xi+400. (London: Chapman and Hall, Ltd., 1945.) 36*s*. net.

X-ray diffraction is now used extensively for investigating the structure of all classes of solids. In metallurgy particularly it has led to results of considerable practical interest. This book summarizes a great deal of material which hitherto has been available only in original scattered papers, and presents a timely and logical account of modern x-ray practice and its contributions to the study of metals and allied substances.

The first part of the book deals with the theory of x-ray diffraction and the practical methods of recording the diffraction patterns. The scope may be indicated by the chapter headings: X-ray generating apparatus, Space lattice, Diffraction of x rays by the crystal lattice, Experimental methods of obtaining diffraction patterns, Influence of the atomic pattern on the intensities of the x-ray reflections, and Crystal structures of the metals. The second part of the book then proceeds to the applications. The main topics, again described by the chapter headings, are: Study of thermal equilibrium diagrams, Measurement of grain size (including effects of deformation), Grain orientation, and Study of refractory materials. The book is then rounded off by a chapter on the older use of x rays in the radiography of metals.

In considering the theory, the author has confined himself to the bare essentials. He has taken the view, quite properly, that the application of x rays to the problems of metallurgy can be fairly appreciated without delving into the mysteries of Fourier synthesis and the theory of space groups; at the same time, it might be remarked in passing that such studies will not impede the x-ray metallographer in his career. Treatment of the practical techniques also is general, except for the Bradley powder-camera, which is described in outstanding detail. It is perhaps to be recommended that a student should first familiarize himself thoroughly with one specific X-ray technique ; and it might perhaps be added that if he can obtain the highest accuracy with this powder camera, he can proceed with confidence to any other method. However, for the practical metallurgist the reduction of a metal to powder, as a preliminary to X-ray examination, puts an unnecessary curb on the versatility of the X-ray method. A fuller description of a camera which will deal equally well with solid specimens of metal might be included with advantage in future editions of the book, which it is hoped will follow in due course.

The chapters on practical applications will be found very useful. The author has not confined himself to the earlier use of X rays in determining atomic arrangements ; he has also given considerable space to the later applications concerned with the macrostructure of metals. Knowledge of the atomic structure is, of course, an essential background to any study of metals, but the atomic bond in metals is not the vital link that it is, for instance, in organic materials. In practice, much coarser features of structure are often of overriding importance ; the metallurgist is often less concerned with the basic atomic arrangement than with larger scale factors such as the segregation of alloying elements and impurities, the dispersion of precipitates, crystallite size and orientation, internal strains, and grain boundary conditions. That is why the X rays will never supplant the microscope in metallurgical research, although, if intelligently used, they will supplement it with considerable advantage. The topics treated in this book well illustrate the range of application of X rays in this field. The metallurgist may consider some points over-emphasized, as the use perhaps of X rays in estimating coarse grain sizes, and some points under-stressed ; but he will undoubtedly find that this part of the book makes interesting and suggestive reading.

The book fills its purpose in providing a guide to the subject, and, at the same time, a handy source of reference. It will be found useful by the student and general reader who wishes rapidly to see the high-lights of the subject, and to the research worker who may wish to refresh himself upon the many practical details involved in the interpretation of the diffraction patterns. A slight drawback is that the book is produced under war economy standards, and the reproduction of photographs, including some from papers by the reviewer, suffers somewhat in comparison with pre-war standards. Perhaps this may soon be remedied.

W. A. W.

INDEX

PAGE
Aberration of a lens system, A differential method of adjusting the 302
Absorption, The influence of, on the shapes and positions of lines in Debye-Scherrer powder photographs 108
Achromatized plate-mirror systems 199
Andrade, E. N. da C., F.R.S. : President of the Physical Society . . . *frontispiece*
Andrade, E. N. da C. : Presidential address, The history and future of the Physical Society xxi
Aplanats, Flat-fielded singlet 567
Area of a search coil, Measurement of the effective 238
Aston, F. W., F.R.S. . . . . *facing* 259

Balance, precision, Recent improvements in a 136
Ballard, W. E. ; The formation of metal-sprayed deposits ; and discussion . 67, 242
Band spectra of tin and lead halides 37
Band spectrum of MgO 12
Barkla, Charles Glover (Obituary notice) 249
Barrow, R. F. and Crawford, D. V. : A note on the spectrum of MgO . . 12
Bell, R. P. : A problem of heat conduction with spherical symmetry . . . 45
Binnie, A. M. : A double-refraction method of detecting turbulence in liquids . 390
Bishop, R. F., Hill, R. and Mott, N. F. : The theory of indentation and hardness tests 147
Brightness and temperature of a total radiator 440
Brightness of present-day dyes ; and corrigendum 15, 146
Brown, F. G. : Exact addition formulae for the axial spherical aberration and curvature of field of an optical system of centred spherical surfaces . . 403
Burch, C. R. : Flat-fielded singlet aplanats 567

Caldin, E. F. : The relation between the brightness and temperature of a total radiator 440
Charlesby, A. : Effect of temperature on the structure of highly polymerized hydrocarbons 510
Charlesby, A. : Structure and orientation in thin films of polythene . . . 496
Choong, S. P. : Coloration and luminescence produced by radium rays in the different varieties of quartz, and some optic properties of these varieties . 49
Catterson-Smith, John Keats (Obituary notice) 254
Coloration and luminescence produced by radium rays in quartz 49
Colorimetry and the colours of total radiators, Illuminants for ; and corrigenda 222, 370
Colour reception, The electrophysiological analysis of the fundamental problem of . 447
Combinations of spherical lenses to replace non-spherical refracting surfaces in optical systems 84
Compton profile for atoms from H to Ne, Theoretical shape of the . . . 190
Corona, A physical theory of the solar 271
Cosmic rays, The geophysical aspect of 464
Coulson, C. A., *see* Duncanson, W. E.
Craven, E. C. : A study of the comparative method of determining gaseous refractivities 97
Crawford, D. V., *see* Barrow, R. F.
Critical potential experiment, A Franck and Hertz : demonstration . . . 577
Cruickshank, F. D. : A system of transfer coefficients for use in the design of lens systems : I–V. 350, 362, 419, 426, 430
Cruickshank, F. D., *see* M'Aulay, A. L.
Crystal lattices, Imperfections of, as investigated by the study of x-ray diffuse scattering 310
Crystals, Extrapolation methods in the derivation of accurate unit-cell dimensions of 160
Cuckow, F. W. : The electron microscope and its applications 564
Curvature of field and spherical aberration of an optical system of centred spherical surfaces 403

PAGE

Davies, C. N. : Definitive equations for the fluid resistance of spheres 259
Debye-Scherrer method, The determination of lattice parameters by the 126
Debye-Scherrer powder photographs, The influence of absorption on the shapes and positions of lines in 108
Deposits, metal-sprayed, The formation of 67, 242
Diatomic molecules, Resonance in precessional states of 32
Differential method of adjusting the aberration of a lens system 302
Diffusion in the ionosphere 519
Dimensions of crystals, Extrapolation methods in the derivation of 160
Directional loci in a magnetic field, and the locating of neutral points 294
Double-refraction method of detecting turbulence in liquids 390
Duddell Medallist, Twenty-first, F. W. Aston, F.R.S. *facing* 259
Duncanson, W. E. and Coulson, C. A. : Theoretical shape of the Compton profile for atoms from H to Ne 190
Duperier, A. : The geophysical aspect of cosmic rays 464
Dyes, The brightness of ; and corrigendum 15, 146
Dynamic measurement of Young's modulus for short wires 412

Eddington, Sir Arthur (Obituary notice) 244
Electron microscope and its applications 564
Electron microscopes (Discussion) 63
Electrophysiological analysis of the fundamental problem of colour reception 447
Equations, Definitive, for the fluid resistance of spheres 259
Expansion, Thermal, of graphite : I, Experimental ; II, Theoretical 477, 486
Extrapolation methods in the derivation of accurate unit-cell dimensions of crystals 160

Films of polythene, Structure and orientation in thin 496
Flat-fielded singlet aplanats 567
Fleming, John Ambrose (Obituary notice) 581
Fluid resistance of spheres, Definitive equations for the 259
Formulae, Exact addition, for the axial spherical aberration and curvature of field of an optical system of centred spherical surfaces 403
Franck and Hertz critical-potential experiment : demonstration 577
Frictionless state of aggregation 371

Gaseous refractivities, The comparative method of determining 97
Gent, W. L., *see* Thomas, A. M.
Geophysical aspect of cosmic rays 464
Goddard, L. S. : Discussion on electron microscopes 63
Goddard, L. S. and Klemperer, O. : Discussion on electron microscopes 63
Granit, Ragnar : The electrophysiological analysis of the fundamental problem of colour reception 447
Graphite, The thermal expansion of : I, Experimental ; II, Theoretical 477, 486
Grüneisen's equation for thermal expansion 209
Guinier, A. J. : Imperfections of crystal lattices as investigated by the study of x-ray diffuse scattering 310
Guthrie Lecture, Twenty-ninth, by A. Duperier 464

H to Ne, Shape of the Compton profile for atoms from 190
Harding, H. G. W. : Illuminants for colorimetry and the colours of total radiators ; and corrigenda 222, 370
Hardness and indentation tests, The theory of 147
Heat conduction with spherical symmetry 45
Hill, R., *see* Bishop, R. F.
Houghton, J. L. : Combination of spherical lenses to replace non-spherical refracting surfaces in optical systems 84
Howell, H. G. : Resonance in precessional states of diatomic molecules 32
Howell, H. G. : The spectra of tin and lead hydrides 37
Hume-Rothery, W. : Grüneisen's equation for thermal expansion 209

PAGE

Hydrides, tin and lead, Spectra of . . . 37
Hydrocarbons, Effect of temperature on the structure of highly polymerized . . . 510

Illuminants for colorimetry and the colours of total radiators ; and corrigenda 222, 370
Imperfections of crystal lattices as investigated by the study of x-ray diffuse scattering . . . 310
Improvements in a precision balance and the efficacy of rhodium plating for standard weights . . . 136
Indentation and hardness tests, The theory of . . . 147
Interferometer, Apparatus illustrating the Michelson stellar . . . 579
Ionosphere, Note on diffusion in the . . . 519

Jaeger, J. C. : Note on diffusion in the ionosphere . . . 519

Klemperer, O., *see* Goddard, L. S.

Lattice parameters, The Debye-Scherrer method of determining . . . 126
Lead and tin hydrides, The spectra of . . . 37
Lens systems, A differential method of adjusting the aberration of a . . . 302
Lens systems, A system of transfer coefficients for use in the design of : I–V. 350, 362, 419, 426, 430
Lenses, Combinations of spherical, to replace non-spherical refracting surfaces in optical systems . . . 84
Lenses, Flat-fielded singlet aplanatic . . . 567
Lenses, Theoretical investigation on telephoto . . . 543
Linfoot, E. H. : Achromatized plate-mirror systems . . . 199
Liquids, A double-refraction method of detecting turbulence in . . . 390
Luminescence and coloration produced by radium rays in quartz . . . 49

M'Aulay, A. L. : A transfer method for deriving the effect on the image formed by an optical system from ray changes produced at a given surface . . . 435
M'Aulay, A. L. and Cruickshank, F. D. : A differential method of adjusting the aberration of a lens system . . . 302
Magnetic field, Directional loci and the locating of neutral points in a . . . 294
Manley, J. J. : Recent improvements in a precision balance and the efficacy of rhodium plating for standard weights . . . 136
Mean scotopic visibility curve : corrigendum . . . 66
Mechanical vibrations of known amplitude, A method for obtaining small . . . 534
Mendelssohn, K. : The frictionless state of aggregation . . . 371
Metal-sprayed deposits, The formation of . . . 67, 242
MgO, Spectrum of . . . 12
Michelson stellar interferometer, Apparatus illustrating the . . . 579
Mikhail, H. and Yousef, Y. L. : The measurement of the effective area of a search coil . . . 238
Mott, N. F., *see* Bishop, R. F.

Nelson, J. B. and Riley, D. P. : An experimental investigation of extrapolation methods in the derivation of accurate unit-cell dimensions of crystals . . . 160
Nelson, J. B. and Riley, D. P. : The thermal expansion of graphite from 15° c. to 800° c. : I. Experimental . . . 477
Neutral points and directional loci in a magnetic field . . . 294
Non-reflecting termination of a transmission line . . . 90

Obituary notices . . . 244, 581
Optic properties in the different varieties of quartz . . . 49
Optical system, A transfer method for deriving the effect on the image formed by an, from ray changes produced at a given surface . . . 435
Optical system of centred spherical surfaces, The axial spherical aberration and curvature of field of an . . . 403

PAGE
Optical system, A differential method of adjusting the aberration of an . . . 302
Optical system, On tracing rays through an . . . 286
Optical systems, Achromatized plate-mirror . . . 199
Optical systems, A system of transfer coefficients for use in the design of ; I–V. 350, 362, 419, 426, 430
Optical systems, Combinations of spherical lenses to replace non-spherical refracting surfaces in . . . 84
Optics, Variational formulae in . . . 558
Owen, D. : Directional loci in a magnetic field, and the locating of neutral points . 294

Pelc, S. R. : The photographic action of x rays . . . 523
Permeation and sorption of water vapour in varnish films . . . 324
Perry, W. E. : Wall- and salt-absorption corrections in radium-content measurements 178
Photocell, The internal resistance of the selenium rectifier . . . 1
Photographic action of x rays . . . 523
Physical Society, The history and future of the . . . xxi
Pincherle, L. : The polarizing angle for reflection at the boundary between two absorbing media . . . 56
Polarizing angle for reflection at the boundary between two absorbing media . 56
Polymerized hydrocarbons, Effect of temperature on the structure of highly . . 510
Polythene, Structure and orientation in thin films of . . . 496
Presidential Address, by E. N. da C. Andrade . . . xxi
Preston, J. S. and Smith, G. W. G. : The internal resistance of the selenium rectifier photocell, with special reference to the sputtered metal film . . 1

Quartz, Coloration and luminescence produced by radium rays in, and some optic properties of . . . 49

Radium-content measurements, Wall- and salt-absorption corrections in . . 178
Ray-tracing through an optical system . . . 286
Recent reports and catalogues . . . 66, 146, 258
Rectangular voltage waves from a low impedance source . . . 60
Reflection at the boundary between two absorbing media, The polarizing angle for . 56
Refractivities, gaseous, The comparative method of determining . . . 97
Rehfisch, T. J. : Rectangular voltage waves from a low impedance source . . 60
Resistance, fluid, of spheres, Definitive equations for . . . 259
Resistance of the selenium rectifier photocell . . . 1
Resonance in precessional states of diatomic molecules . . . 32
Reviews of books . . . 145, 255, 368, 444, 586
Rhodium plating for standard weights, The efficacy of . . . 136
Riley, D. P., *see* Nelson, J. B.
Riley, D. P. : The thermal expansion of graphite : II. Theoretical . . . 486

Saha, Meghnad : A physical theory of the solar corona . . . 271
Sanders, J. H. : Demonstration : A Franck and Hertz critical-potential experiment 577
Scattering, X-ray diffuse ; imperfections of crystal lattices as investigated by the study of . . . 310
Scotopic visibility curve : corrigendum . . . 66
Search coil, The measurement of the effective area of a . . . 238
Selenium rectifier photocell, The internal resistance of the . . . 1
Sinclair, H., *see* Taylor, A.
Skinner, Sidney (Obituary notice) . . . 584
Smith, D. H. : A method for obtaining small mechanical vibrations of known amplitude . . . 534
Smith, D. H. : The non-reflecting termination of a transmission line . . . 90
Smith, G. W. G., *see* Preston, J. S.
Smith, John Roderick Ennis (Obituary notice) . . . 253
Smith, T. : On tracing rays through an optical system . . . 286
Smith, T., *see* Stiles, W. S.
Smith, T. : Theoretical investigation on telephoto lenses . . . 543

PAGE
Smith, T. : Variational formulae in optics . . . 558
Solar corona, A physical theory of the . . . 271
Sorption and permeation of water vapour in varnish films . . . 324
Spectra of tin and lead hydrides . . . 37
Spectrum of MgO . . . 12
Spheres, Definitive equations for the fluid resistance of . . . 259
Spherical aberration and curvature of field of an optical system of centred spherical surfaces . . . 403
Spherical symmetry, Heat conduction with . . . 45
Sprayed metal deposits, The formation of . . . 67, 242
State of aggregation, The frictionless . . . 371
Stellar interferometer, Apparatus illustrating the : demonstration . . . 579
Stiles, W. S. and Smith, T. : A mean scotopic visibility curve : corrigendum . . . 66

Taylor, A. and Sinclair, H. : The determination of lattice parameters by the Debye-Scherrer method . . . 126
Taylor, A. and Sinclair, H. : The influence of absorption on the shapes and positions of lines in Debye-Scherrer powder photographs . . . 108
Telephoto lenses, Theoretical investigation on . . . 543
Temperature and brightness of a total radiator . . . 440
Temperature effect on the structure of highly polymerized hydrocarbons . . . 510
Theory of indentation and hardness tests . . . 147
Theory of the solar corona . . . 271
Thermal expansion of graphite : I, Experimental ; II, Theoretical . . . 477, 486
Thermal expansion, Grüneisen's equation for . . . 209
Thomas, A. M. and Gent, W. L. : Permeation and sorption of water vapour in varnish films . . . 324
Thomas Young Oration, Fourteenth, by Ragnar Granit . . . 447
Tin and lead hydrides, The spectra of . . . 37
Total radiator, the relation between The brightness and temperature of a . . . 440
Total radiators, Illuminants for colorimetry and the colours of ; and corrigenda 222, 370
Transfer coefficients for use in the design of lens systems ; I–V . 350, 362, 419, 426, 430
Transfer method for deriving the effect on the image formed by an optical system from ray changes produced at a given surface . . . 435
Transmission line, The non-reflecting termination of a . . . 90
Treanor, P. J. : Apparatus illustrating the Michelson stellar interferometer : demonstration . . . 579
Tscherning, Marius Hans Erik (Obituary notice) . . . 253
Turbulence in liquids, A double-refraction method of detecting . . . 390

Unit-cell dimensions of crystals, Extrapolation methods in the accurate derivation of 160

Varnish films, Permeation and sorption of water vapour in . . . 324
Variational formulae in optics . . . 558
Vibrations of known amplitude, A method for obtaining small mechanical . . . 534
Vickerstaff, T. : The brightness of present-day dyes ; and corrigendum . . . 15, 146
Visibility curve, Mean scotopic : corrigendum . . . 66

Wall- and salt-absorption corrections in radium-content measurements . . . 178
Waves from a low impedance source, Rectangular voltage . . . 60
Weights, The efficacy of rhodium plating for standard . . . 136
Wilberforce, L. R. (Obituary notice) . . . 585

X-ray diffuse scattering, Imperfections of crystal lattices as investigated by the study of . . . 310
X rays, The photographic action of . . . 523

Young, James (Obituary notice) . . . 585
Young's modulus for short wires, Dynamic measurement of . . . 412
Yousef, Y. L. : Dynamic measurement of Young's modulus for short wires . . . 412
Yousef, Y. L., *see* Mikhail, H.

INDEX TO REVIEWS OF BOOKS

PAGE

Acoustics Committee of the Building Research Board of the Department of Scientific and Industrial Research : *Sound Insulation and Acoustics* 258

Brandenburg, H. : *Sechsstellige trigonometrische Tafeln* 370, 446

Carter, G. W. : *The Simple Calculation of Electrical Transients* 145
Creasey, C. H., *see* Eve, A. S.

Dorsey, N. Ernest : *The Velocity of Light* 368

Emde, F. : *Tables of Elementary Functions* 368, 446
Eve, A. S. and Creasey, C. H. : *Life and Work of John Tyndall* 588

Granier, J. : *Introduction a l'étude des champs physiques* 445

Heitler, W. : *The Quantum Theory of Radiation* 446
His Majesty's Stationery Office : *Five-figure Logarithm Tables* 145

Industrial Radiology Group of the Institute of Physics : *Handbook of Industrial Radiology* 257

Johnson, B. K. : *Practical Optics* 586

Lyons, Sir Henry : *The Royal Society*, 1660–1940 : *A History of its Administration under the Charters* 255

Manley, R. G. : *Waveform Analysis* 444

Rocard, Yves : *Théorie des Oscillateurs* 587

Schrödinger, Erwin : *What is Life ?* 146

Taylor, A. : *An Introduction to X-ray Metallurgy* 589

Wright, W. D. : *The Measurement of Colour* 368

Yarwood, T. M. : *School Physics* 370

157